

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 18, 2003, 19:53:48 ; Search time 4471.78 Seconds  
(without alignments)  
9743.047 Million cell updates/sec

Title: US-09-987-455-7  
Perfect score: 1065  
Sequence: 1 tctgagggaacagtgactg.....gtgacaacatgcgaccgtga 1065

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 20454813386 residues  
Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl.\*

- 1: gb\_ba.\*
- 2: gb\_htg.\*
- 3: gb\_in.\*
- 4: gb\_om.\*
- 5: gb\_ov.\*
- 6: gb\_pat.\*
- 7: gb\_ph.\*
- 8: gb\_pl.\*
- 9: gb\_pr.\*
- 10: gb\_ro.\*
- 11: gb\_sts.\*
- 12: gb\_sy.\*
- 13: gb\_un.\*
- 14: gb\_vi.\*
- 15: em\_ba.\*
- 16: em\_fun.\*
- 17: em\_hum.\*
- 18: em\_in.\*
- 19: em\_mu.\*
- 20: em\_om.\*
- 21: em\_or.\*
- 22: em\_ov.\*
- 23: em\_pat.\*
- 24: em\_ph.\*
- 25: em\_pl.\*
- 26: em\_ro.\*
- 27: em\_sts.\*
- 28: em\_un.\*
- 29: em\_vi.\*
- 30: em\_htg\_hum.\*
- 31: em\_htg\_inv.\*
- 32: em\_htg\_other.\*
- 33: em\_htg\_mus.\*
- 34: em\_htg\_pln.\*
- 35: em\_htg\_rtd.\*
- 36: em\_htg\_mam.\*
- 37: em\_htg\_vrt.\*
- 38: em\_sy.\*
- 39: em\_htgo\_hum.\*
- 40: em\_htgo\_mus.\*
- 41: em\_htgo\_other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1065	100.0	1065	6	AX431270	AX431270 Sequence
2	1065	100.0	1065	6	AX431273	AX431273 Sequence
3	1065	100.0	1068	6	A27431	A27431 DNA sequence
4	1065	100.0	1068	6	AR059987	AR059987 Sequence
5	1065	100.0	1068	6	E01935	E01935 Synthetic D
6	1065	100.0	1170	6	A27727	A27727 DNA sequenc
7	1065	100.0	1170	6	E01945	E01945 Synthetic D
8	1065	100.0	1314	6	A27435	A27435 DNA sequenc
9	1065	100.0	1314	6	E01937	E01937 Synthetic D
10	1065	100.0	1419	6	A27725	A27725 DNA sequenc
11	1065	100.0	1419	6	AR059996	AR059996 Sequence
12	1065	100.0	1419	6	E01944	E01944 Synthetic D
13	1065	100.0	1689	6	I06609	I06609 Sequence 44
14	1065	100.0	1689	6	I06614	I06614 Sequence 46
15	1065	100.0	1689	6	I08789	I08789 Sequence 3
16	1065	100.0	1780	6	A06611	A06611 Synthetic n
17	1065	100.0	1780	6	A10226	A10226 Synthetic D
18	1065	100.0	1836	6	E01176	E01176 DNA encodin
19	1065	100.0	1836	6	I01581	I01581 Sequence 1
20	1065	100.0	1981	6	I01583	I01583 Sequence 3
21	1065	100.0	1989	6	I01256	I01256 Sequence 2
22	1065	100.0	2099	6	A07207	A07207 Artificial
23	1065	100.0	2100	6	A07197	A07197 H.sapiens m
24	1065	100.0	2100	6	A27429	A27429 cDNA sequen
25	1065	100.0	2100	6	E01934	E01934 DNA encodin
26	1065	100.0	2101	6	AR059986	AR059986 Sequence
27	1065	100.0	2151	6	E00654	E00654 cDNA encodi
28	1065	100.0	2162	6	AR017909	AR017909 Sequence
29	1065	100.0	2162	6	AR017931	AR017931 Sequence
30	1065	100.0	2162	6	I18796	I18796 Sequence 3
31	1065	100.0	2249	6	A07209	A07209 Artificial
32	1065	100.0	2544	6	I08530	I08530 Sequence 1
33	1065	100.0	2547	6	E01163	E01163 cDNA encodi
34	1065	100.0	2547	6	I07841	I07841 Sequence 1
35	1065	100.0	2547	6	I07989	I07989 Sequence 1
36	1065	100.0	2547	6	I09622	I09622 Sequence 1
37	1065	100.0	2549	6	E01055	E01055 DNA encodin
38	1065	100.0	2560	6	A01465	A01465 H.sapiens t
39	1065	100.0	7533	6	I08638	I08638 Sequence 2
40	1064	99.9	1689	6	I08637	I08637 Sequence 1
41	1063.4	99.8	1170	6	AR059997	AR059997 Sequence
42	1063.4	99.8	1314	6	AR059989	AR059989 Sequence
43	1063.4	99.8	1506	6	E02247	E02247 cDNA sequen
44	1063.4	99.8	1680	6	I06422	I06422 Sequence 40
45	1063.4	99.8	1689	6	A14927	A14927 tPA-DNA fro

ALIGNMENTS

RESULT 1  
AX431270  
LOCUS AX431270 1065 bp DNA linear PAT 28-JUN-2002  
DEFINITION Sequence 4 from Patent WO240650.  
ACCESSION AX431270  
VERSION AX431270.1 GI:21656152  
KEYWORDS synthetic construct  
SOURCE synthetic construct  
ORGANISM synthetic construct  
REFERENCE 1  
Goetz,F., Werner,R.G., Tayapiwatana,C., Manosroi,A. and Manosroi,J.  
METHODS for large scale production of recombinant dna-derived tpa  
or k2s molecules  
TITLE Patent: WO 0240650-A 4 23-MAY-2002;  
JOURNAL

```
BOEHRINGER INGELHEIM INT (DE)
FEATURES
  Location/Qualifiers
    1..1065
    /organism="synthetic construct"
    /mol_type="synthetic construct"
    /db_xref="taxon:32630"
    /note="coding sequence for K2S protein"
BASE COUNT      225 a   314 c   312 g   214 t
ORIGIN
Query Match      100.0%; Score 1065; DB 6; Length 1065;
Best Local Similarity 100.0%; Pred. No. 4.7e-248;
Matches 1065; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCTGAGGAAACAGTACTGCTTCTGGAATGGTCACTACCGTGGCAGCACAGC 60
DB 1 TCTGAGGAAACAGTACTGCTTCTGGAATGGTCACTACCGTGGCAGCACAGC 60
QY 61 CTCACCGAGTGGGTCCTCTGCTCCCTCCGTTGGAAATTCATGATCCTGATAGGCAAGTT 120
DB 61 CTCACCGAGTGGGTCCTCTGCTCCCTCCGTTGGAAATTCATGATCCTGATAGGCAAGTT 120
QY 121 TACACAGACAGAACCCAGTGGCCAGGACCTGGGCTGGGCAACATAATTAATCTGCCGG 180
DB 121 TACACAGACAGAACCCAGTGGCCAGGACCTGGGCTGGGCAACATAATTAATCTGCCGG 180
QY 181 AATCTGTATGGGATGCCAGCCCTGGTGGCCAGCTGGGCTGGGCAACATAATTAATCTGCCGG 240
DB 181 AATCTGTATGGGATGCCAGCCCTGGTGGCCAGCTGGGCTGGGCAACATAATTAATCTGCCGG 240
QY 241 GAGTACTGTGATGTCCTCTGCTCCCTCCGTTGGAAATTCATGATCCTGATAGGCAAGTT 300
DB 241 GAGTACTGTGATGTCCTCTGCTCCCTCCGTTGGAAATTCATGATCCTGATAGGCAAGTT 300
QY 301 TTTCGCATCAAGAGAGGCTCTTCCGACAGATCGCTCCACCTCGGAGGCTGGCCTATC 360
DB 301 TTTCGCATCAAGAGAGGCTCTTCCGACAGATCGCTCCACCTCGGAGGCTGGCCTATC 360
QY 361 TTTCGCAAGCAGAGAGGCTCGCCGAGAGCGGTTCTGTGGGGGATATCTATCAGC 420
DB 361 TTTCGCAAGCAGAGAGGCTCGCCGAGAGCGGTTCTGTGGGGGATATCTATCAGC 420
QY 421 TCCTGCTGGATTCCTCTGCGGCCCACTGCTTCCAGGAGAGGTTTCCGCCCCACCCACTG 480
DB 421 TCCTGCTGGATTCCTCTGCGGCCCACTGCTTCCAGGAGAGGTTTCCGCCCCACCCACTG 480
QY 481 ACGGTGATCTGGGAGAGACATACCGGGTGGTCCCTGGCGAGGAGGAGCAAAATTTGAA 540
DB 481 ACGGTGATCTGGGAGAGACATACCGGGTGGTCCCTGGCGAGGAGGAGCAAAATTTGAA 540
QY 541 GTCGAAAATACATTGTCCATAAGGAATTCCGATGATGACATTACGACAATGACATTGCG 600
DB 541 GTCGAAAATACATTGTCCATAAGGAATTCCGATGATGACATTACGACAATGACATTGCG 600
QY 601 CTGCTGCAGCTGAATTCGGATTCGTCCTGCTGTCCTCCAGGAGAGCGTGGTCCGCACT 660
DB 601 CTGCTGCAGCTGAATTCGGATTCGTCCTGCTGTCCTCCAGGAGAGCGTGGTCCGCACT 660
QY 661 GTGTGCTTCCCGCGGAGACCTGAGCTGCGGACTGGAGCTGGAGCTCTCGGC 720
DB 661 GTGTGCTTCCCGCGGAGACCTGAGCTGCGGACTGGAGCTGGAGCTCTCGGC 720
QY 721 TACGCAAGCATGAGGCTTGTCTCTTCTATTTCGGAGCGGCTGAAGAGGCTCATGTC 780
DB 721 TACGCAAGCATGAGGCTTGTCTCTTCTATTTCGGAGCGGCTGAAGAGGCTCATGTC 780
QY 781 AGACTGTACCCATCCAGCGCTGCAATCAACAATTTACTTAACAGAAACAGTCAACCGAC 840
DB 781 AGACTGTACCCATCCAGCGCTGCAATCAACAATTTACTTAACAGAAACAGTCAACCGAC 840
QY 841 AACATGCTGTGTGGAGACACTCGGAGCGGCGGCGCCAGGCAAACTTGCACGAGCC 900
DB 841 AACATGCTGTGTGGAGACACTCGGAGCGGCGGCGCCAGGCAAACTTGCACGAGCC 900
Dk
```

```

QY 901 TGCCAGGCGGATTCGGAGGCCCCCTGCTGTGTCTGAACGATGGCCGATGACTTTGGTG 960
DB 901 TGCCAGGCGGATTCGGAGGCCCCCTGCTGTGTCTGAACGATGGCCGATGACTTTGGTG 960
QY 961 GGCAATCATAGCTGGGCGCTGGGCTGTGGACAGAAAGATGTCGCGGTGTGTACACAAAG 1020
DB 961 GGCAATCATAGCTGGGCGCTGGGCTGTGGACAGAAAGATGTCGCGGTGTGTACACAAAG 1020
QY 1021 GTTACCAACTACCTAGACTGGATTCTGTGACACATGCGACCGTGA 1065
DB 1021 GTTACCAACTACCTAGACTGGATTCTGTGACACATGCGACCGTGA 1065

RESULT 2
AX431273 LOCUS      1065 bp      DNA      linear      PAT 28-JUN-2002
DEFINITION Sequence 7 from Patent WO0240650.
ACCESSION  AX431273
VERSION     AX431273.1  GI:21656155
KEYWORDS   . synthetic construct
SOURCE      synthetic construct
ORGANISM    artificial sequences.
REFERENCE   1
AUTHORS     Goetz, F., Werner, R.G., Tayapiwatana, C., Manosroi, A. and Manosroi, J.
TITLE       Methods for large scale production of recombinant dna-derived tpa
JOURNAL     or k2s molecules
PUBLISHED  Patent: WO 0240650-A 7 23-MAY-2002;
BOEHRINGER INGELHEIM INT (DE)
FEATURES
  Location/Qualifiers
    1..1065
    /organism="synthetic construct"
    /mol_type="genomic DNA"
    /db_xref="taxon:32630"
    /note="coding sequence for K2S protein"
BASE COUNT      225 a   314 c   312 g   214 t
ORIGIN
Query Match      100.0%; Score 1065; DB 6; Length 1065;
Best Local Similarity 100.0%; Pred. No. 4.7e-248;
Matches 1065; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCTGAGGAAACAGTACTGCTTCTGGAATGGTCACTACCGTGGCAGCACAGC 60
DB 1 TCTGAGGAAACAGTACTGCTTCTGGAATGGTCACTACCGTGGCAGCACAGC 60
QY 61 CTCACCGAGTGGGTCCTCTGCTCCCTCCGTTGGAAATTCATGATCCTGATAGGCAAGTT 120
DB 61 CTCACCGAGTGGGTCCTCTGCTCCCTCCGTTGGAAATTCATGATCCTGATAGGCAAGTT 120
QY 121 TACACAGACAGAACCCAGTGGCCAGGACCTGGGCTGGGCAACATAATTAATCTGCCGG 180
DB 121 TACACAGACAGAACCCAGTGGCCAGGACCTGGGCTGGGCAACATAATTAATCTGCCGG 180
QY 181 AATCTGTATGGGATGCCAGCCCTGGTGGCCAGCTGGGCTGGGCAACATAATTAATCTGCCGG 240
DB 181 AATCTGTATGGGATGCCAGCCCTGGTGGCCAGCTGGGCTGGGCAACATAATTAATCTGCCGG 240
QY 241 GAGTACTGTGATGTCCTCTGCTCCCTCCGTTGGAAATTCATGATCCTGATAGGCAAGTT 300
DB 241 GAGTACTGTGATGTCCTCTGCTCCCTCCGTTGGAAATTCATGATCCTGATAGGCAAGTT 300
QY 301 TTTCGCATCAAGAGAGGCTCTTCCGACAGATCGCTCCACCTCGGAGGCTGGCCTATC 360
DB 301 TTTCGCATCAAGAGAGGCTCTTCCGACAGATCGCTCCACCTCGGAGGCTGGCCTATC 360
QY 361 TTTCGCAAGCAGAGAGGCTCGCCGAGAGCGGTTCTGTGGGGGATATCTATCAGC 420
DB 361 TTTCGCAAGCAGAGAGGCTCGCCGAGAGCGGTTCTGTGGGGGATATCTATCAGC 420
QY 421 TCCTGCTGGATTCCTCTGCGGCCCACTGCTTCCAGGAGAGGTTTCCGCCCCACCCACTG 480
DB 421 TCCTGCTGGATTCCTCTGCGGCCCACTGCTTCCAGGAGAGGTTTCCGCCCCACCCACTG 480
QY 481 ACGGTGATCTGGGAGAGACATACCGGGTGGTCCCTGGCGAGGAGGAGCAAAATTTGAA 540
DB 481 ACGGTGATCTGGGAGAGACATACCGGGTGGTCCCTGGCGAGGAGGAGCAAAATTTGAA 540
QY 541 GTCGAAAATACATTGTCCATAAGGAATTCCGATGATGACATTACGACAATGACATTGCG 600
DB 541 GTCGAAAATACATTGTCCATAAGGAATTCCGATGATGACATTACGACAATGACATTGCG 600
QY 601 CTGCTGCAGCTGAATTCGGATTCGTCCTGCTGTCCTCCAGGAGAGCGTGGTCCGCACT 660
DB 601 CTGCTGCAGCTGAATTCGGATTCGTCCTGCTGTCCTCCAGGAGAGCGTGGTCCGCACT 660
QY 661 GTGTGCTTCCCGCGGAGACCTGAGCTGCGGACTGGAGCTGGAGCTCTCGGC 720
DB 661 GTGTGCTTCCCGCGGAGACCTGAGCTGCGGACTGGAGCTGGAGCTCTCGGC 720
QY 721 TACGCAAGCATGAGGCTTGTCTCTTCTATTTCGGAGCGGCTGAAGAGGCTCATGTC 780
DB 721 TACGCAAGCATGAGGCTTGTCTCTTCTATTTCGGAGCGGCTGAAGAGGCTCATGTC 780
QY 781 AGACTGTACCCATCCAGCGCTGCAATCAACAATTTACTTAACAGAAACAGTCAACCGAC 840
DB 781 AGACTGTACCCATCCAGCGCTGCAATCAACAATTTACTTAACAGAAACAGTCAACCGAC 840
QY 841 AACATGCTGTGTGGAGACACTCGGAGCGGCGGCGCCAGGCAAACTTGCACGAGCC 900
DB 841 AACATGCTGTGTGGAGACACTCGGAGCGGCGGCGCCAGGCAAACTTGCACGAGCC 900
```



```
Db 421 TCTGCTGGATTCTCTGCGGCCCACTGCTTTCAGGAGAGGTTTCCGCCCCACACACCTG 480
Qy 481 ACGGTGATCTTGGGAGAACATACACGGGTGGTCCCTGCGAGAGGAGAGAGAAATTGAA 540
Db 481 ACGGTGATCTTGGGAGAACATACACGGGTGGTCCCTGCGAGAGGAGAGAGAAATTGAA 540
Qy 541 GTCGAAATATACATTGTCCATAGGAATTCGATGATGACACTTACGACAAATGACATTGG 600
Db 541 GTCGAAATATACATTGTCCATAGGAATTCGATGATGACACTTACGACAAATGACATTGG 600
Qy 601 CTGCTGACGCTGAATTCGATTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
Db 601 CTGCTGACGCTGAATTCGATTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
Qy 661 GTGTGCTTCCCGCGGACCTGACGCTGCGGACTGACGAGTGTGAGCTCTCCGCG 720
Db 661 GTGTGCTTCCCGCGGACCTGACGCTGCGGACTGACGAGTGTGAGCTCTCCGCG 720
Qy 721 TACGGCAAGCATGAGGCTTGTCTCTTCTATTCGAGCGGCTGAAGAGGCTCATGTC 780
Db 721 TACGGCAAGCATGAGGCTTGTCTCTTCTATTCGAGCGGCTGAAGAGGCTCATGTC 780
Qy 781 AGACTGTACCATCCAGCGCTGCACATCAACAATTTACTTAAACAGAACTGACCGAC 840
Db 781 AGACTGTACCATCCAGCGCTGCACATCAACAATTTACTTAAACAGAACTGACCGAC 840
Qy 841 AACATGCTGTGTGCTGGAGACACTCGAGCGCGCGGCCCGGCAAACTTGCACGACCC 900
Db 841 AACATGCTGTGTGCTGGAGACACTCGAGCGCGCGGCCCGGCAAACTTGCACGACCC 900
Qy 901 TGCAGGCGGATTCGGAGGCGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960
Db 901 TGCAGGCGGATTCGGAGGCGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960
Qy 961 GGCATCATCAGCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020
Db 961 GGCATCATCAGCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020
Qy 1021 GTTACCAACTACCTAGACTGATTCGTTGACAACTGCGACCGTGA 1065
Db 1021 GTTACCAACTACCTAGACTGATTCGTTGACAACTGCGACCGTGA 1065

RESULT 3
LOCUS A27431 1068 bp DNA linear PAT 09-JUL-2002
DEFINITION DNA sequence of coding region in pTRPAdeltatrp.
ACCESSION A27431
VERSION A27431.1 GI:21727228
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM artificial sequences
REFERENCE 1 (bases 1 to 1068)
AUTHORS Niwa,M., Saiko,Y., Sasaki,H., Hayashi,M., Notani,J. and Kobayashi,M.
TITLE New tissue plasminogen activator
JOURNAL Patent: EP 0302456-A 31 08-FEB-1989;
FUJISAWA PHARMACEUTICAL CO., LTD
FEATURES
source
1..1068
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
BASE COUNT 226 a 314 c 313 g 215 t
ORIGIN
Query Match 100.0%; Score 1065; DB 6; Length 1068;
Best Local Similarity 100.0%; Pred. No. 4.7e-248;
Matches 1065; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 TCTGAGGAAACAGTACTGCTACTTTGGGAATGGGTACGCTACCGTGGCAACGACAGC 60
|||||
```

RESULT 4  
AR059987

```
Db 4 TCTGAGGAAACAGTACTGCTACTTTTGGGAATGGGTACGCTACCGTGGCAACGACAGC 63
Qy 61 CTCAACGAGTCCGGTGGCTCTCTGCTCCCGTGGAAATTCATGATCCTCATAGGCAAGGTT 120
Db 64 CTCAACGAGTCCGGTGGCTCTCTGCTCCCGTGGAAATTCATGATCCTCATAGGCAAGGTT 123
Qy 121 TACACAGCACAGAACCCAGAGTGGCCAGGCACTGGGCTTGGGCAAAACATAATATTCTGCCGG 180
Db 124 TACACAGCACAGAACCCAGAGTGGCCAGGCACTGGGCTTGGGCAAAACATAATATTCTGCCGG 183
Qy 181 AATCCTGATGGGATGCCAAGCCCTGGTGGCCACGCTGTGAGAACCCGAGCTGACCTGG 240
Db 184 AATCCTGATGGGATGCCAAGCCCTGGTGGCCACGCTGTGAGAACCCGAGCTGACCTGG 243
Qy 241 GAGTACTGTGATGTGCTCTCTGCTCCACCTGCGGCTGAGACAGTACAGCCAGCTCAG 300
Db 244 GAGTACTGTGATGTGCTCTCTGCTCCACCTGCGGCTGAGACAGTACAGCCAGCTCAG 303
Qy 301 TTTCCGATCAAGAGGAGGCTCTTCCGCGACATGCGCTCCACCCCTGGCAGGCTGCCATC 360
Db 304 TTTCCGATCAAGAGGAGGCTCTTCCGCGACATGCGCTCCACCCCTGGCAGGCTGCCATC 363
Qy 361 TTTCCGAGCACAGAGGCTGCGCCGAGAGCGGTTTCTGTGCGGGGCGATACTCATCAGC 420
Db 364 TTTCCGAGCACAGAGGCTGCGCCGAGAGCGGTTTCTGTGCGGGGCGATACTCATCAGC 423
Qy 421 TCCTGTGGATTCTCTCTGCGCCCACTGCTTCCAGGAGAGGTTTCCGCCCCCACCACCTG 480
Db 424 TCCTGTGGATTCTCTCTGCGCCCACTGCTTCCAGGAGAGGTTTCCGCCCCCACCACCTG 483
Qy 481 ACGTGTATCTTGGCGAGAACATACCGGGTGGTCTCTGGCGAGGAGGAGAGAAATTTGAA 540
Db 484 ACGTGTATCTTGGCGAGAACATACCGGGTGGTCTCTGGCGAGGAGGAGAGAAATTTGAA 543
Qy 541 GTCCAAAATATACATTGTGCCATTAAGGAATTCGATGATGACACTTACGACAAATGACATTGCG 600
Db 544 GTCCAAAATATACATTGTGCCATTAAGGAATTCGATGATGACACTTACGACAAATGACATTGCG 603
Qy 601 CTGCTGACGCTGAAATCGGATTCGCTCCGCTGTGCCAGAGAGAGAGCTGGTCCGCACT 660
Db 604 CTGCTGACGCTGAAATCGGATTCGCTCCGCTGTGCCAGGAGAGAGAGCTGGTCCGCACT 663
Qy 661 GTGTGCTTCCCGCGGACCTGACGCTGCGGACTGAGAGTGTGAGCTCTCCGCG 720
Db 664 GTGTGCTTCCCGCGGACCTGACGCTGCGGACTGAGAGTGTGAGCTCTCCGCG 723
Qy 721 TACGGCAAGCATGAGGCTTGTCTCTTCTATTCGAGCGGCTGAGAGGAGCTCATGTC 780
Db 724 TACGGCAAGCATGAGGCTTGTCTCTTCTATTCGAGCGGCTGAGAGGAGCTCATGTC 783
Qy 781 AGACTGTACCATCCAGCGCTGCACATCAACAATTTACTTAAACAGAACTGACCGAC 840
Db 784 AGACTGTACCATCCAGCGCTGCACATCAACAATTTACTTAAACAGAACTGACCGAC 843
Qy 841 AACATGCTGTGTGCTGGAGACACTCGAGCGCGCGGCCCGGCAAACTTGCACGAGCC 900
Db 844 AACATGCTGTGTGCTGGAGACACTCGAGCGCGCGGCCCGGCAAACTTGCACGAGCC 903
Qy 901 TGCAGGCGGATTCGGAGGCGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960
Db 904 TGCAGGCGGATTCGGAGGCGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 963
Qy 961 GGCATCATCAGCTGGGCTTGGGCTGAGAGGATGTCCCGGTGTGTACACAAAG 1020
Db 964 GGCATCATCAGCTGGGCTTGGGCTGAGAGGATGTCCCGGTGTGTACACAAAG 1023
Qy 1021 GTTACCAACTACCTAGACTGATTCGTTGACAACTGCGACCGTGA 1065
Db 1024 GTTACCAACTACCTAGACTGATTCGTTGACAACTGCGACCGTGA 1068
```

LOCUS AR059987 1068 bp DNA linear PAT 29-SEP-1999  
DEFINITION Sequence 44 from patent US 5840533.  
ACCESSION AR059987  
VERSION AR059987.1 GI:5986437  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE Unclassified.  
AUTHORS 1 (bases 1 to 1068)  
Niwa,M., Saito,Y., Sasaki,H., Hayashi,M., Notani,J. and  
Kobayashi,M.  
TITLE Tissue plasminogen activator  
JOURNAL Patent: US 5840533-A 44 24-NOV-1998;  
FEATURES  
Location/Qualifiers  
source 1..1068  
/organism="unknown"  
BASE COUNT 226 a 314 c 313 g 215 t  
ORIGIN  
Query Match 100.0%; Score 1065; DB 6; Length 1068;  
Best Local Similarity 100.0%; Pred. No. 4.7e-248;  
Matches 1065; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 TCTGAGGAAACAGTACTGCTACTTTGGGAATGGGTGAGCTACCGTGGCAACGACAGC 60  
DB 4 TCTGAGGAAACAGTACTGCTACTTTGGGAATGGGTGAGCTACCGTGGCAACGACAGC 63  
QY 61 CTCACGAGTGGGTGCTCTGCTCCGCTGGGAATTCATGATCCTGATAGGCAAGTT 120  
DB 64 CTCACGAGTGGGTGCTCTGCTCCGCTGGGAATTCATGATCCTGATAGGCAAGTT 123  
QY 121 TACACAGACAGAACCCAGTGCCAGGACCTGGGCTGGGCAACATAATTAATCTGCCG 180  
DB 124 TACACAGACAGAACCCAGTGCCAGGACCTGGGCTGGGCAACATAATTAATCTGCCG 183  
QY 181 AATCTGATGGGATGCCAAGCCCTGTGTCACAGCTGCTGAAGAACCCGAGCTGACGTGG 240  
DB 184 AATCTGATGGGATGCCAAGCCCTGTGTCACAGCTGCTGAAGAACCCGAGCTGACGTGG 243  
QY 241 GAGTACTGTGATGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300  
DB 244 GAGTACTGTGATGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 303  
QY 301 TTTCGCATCAAGAGGAGGCTCTTCCGCGACATCGCTCCACCCCTGGCAGGCTGCCATC 360  
DB 304 TTTCGCATCAAGAGGAGGCTCTTCCGCGACATCGCTCCACCCCTGGCAGGCTGCCATC 363  
QY 361 TTTCGCATCAAGAGGAGGCTCTTCCGCGACATCGCTCCACCCCTGGCAGGCTGCCATC 420  
DB 364 TTTCGCATCAAGAGGAGGCTCTTCCGCGACATCGCTCCACCCCTGGCAGGCTGCCATC 423  
QY 421 TCCTGCTGGATCTCTCTGCGCCCACTGCTTCCAGGAGAGGTTTCCGCGCCCAACCTG 480  
DB 424 TCCTGCTGGATCTCTCTGCGCCCACTGCTTCCAGGAGAGGTTTCCGCGCCCAACCTG 483  
QY 481 ACGGTGATCTTGGGAGAACATACCCGCTGGTCTGCTGCGAGGAGGAGCAAGAAATTGAA 540  
DB 484 ACGGTGATCTTGGGAGAACATACCCGCTGGTCTGCTGCGAGGAGGAGCAAGAAATTGAA 543  
QY 541 GTGCAAAATACATTTGTCATTAAGGAATTCGATGATGACCTTACGACAAATGACATTCG 600  
DB 544 GTGCAAAATACATTTGTCATTAAGGAATTCGATGATGACCTTACGACAAATGACATTCG 603  
QY 601 CTGCTGACGCTGAATTCGATTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660  
DB 604 CTGCTGACGCTGAATTCGATTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 663  
QY 661 GTGTGCTTCCCGCGGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720  
DB 664 GTGTGCTTCCCGCGGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 723  
QY 721 TACGCAAGAGATAGGCTTGTCTCTTCTATTCGAGCGGCTGAAGAGGCTCATGTC 780

DB 724 TACGCAAGCATGAGGCTTGTCTCTCTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 783  
QY 781 AGACTGTACCATCCAGCCGCTGCACATCAACATTTACTTACACAGCAGTCAACGAC 840  
DB 784 AGACTGTACCATCCAGCCGCTGCACATCAACATTTACTTACACAGCAGTCAACGAC 843  
QY 841 AACATGCTGTGTCTGGAGACACTCGGAGCGGGGCCCCAGGCAAACTTTCACGAGGCC 900  
DB 844 AACATGCTGTGTCTGGAGACACTCGGAGCGGGGCCCCAGGCAAACTTTCACGAGGCC 903  
QY 901 TGCAGGGCGATTTCGGAGGCCCCCTGCTGTGTCTGAACGATGGCCGCGATGCTTTGGTG 960  
DB 904 TGCAGGGCGATTTCGGAGGCCCCCTGCTGTGTCTGAACGATGGCCGCGATGCTTTGGTG 963  
QY 961 GGCATCATCAGCTGGGCTGGGCTGGGAGAGGATCTCCGGGTGTGTACACAAAG 1020  
DB 964 GGCATCATCAGCTGGGCTGGGCTGGGAGAGGATCTCCGGGTGTGTACACAAAG 1023  
QY 1021 GTTACCAACTACCTAGACTGGATTCTGTGACAAATGCGACCGTGA 1065  
DB 1024 GTTACCAACTACCTAGACTGGATTCTGTGACAAATGCGACCGTGA 1068  
RESULT 5  
E01935  
LOCUS Synthetic DNA encoding new tissue plasminogen activator (t-PA).  
DEFINITION E01935  
ACCESSION E01935  
VERSION E01935.1 GI:2170183  
KEYWORDS JP 1989104167-A/2.  
SOURCE synthetic construct  
ORGANISM synthetic construct  
REFERENCE 1 (bases 1 to 1068)  
AUTHORS Niwa,M., Saito,Y., Sasaki,H., Hayashi,M., Notani,J. and  
Kobayashi,M.  
TITLE NOVEL TISSUE PLASMINOGEN ACTIVATION FACTOR  
JOURNAL Patent: JP 1989104167-A 2 21-APR-1989;  
FUJISAWA PHARMACEUT CO LTD  
COMMENT OS Artificial gene  
OC Artificial sequence; Genes.  
PN JP 1989104167-A/2  
PD 21-APR-1989  
PF 01-AUG-1988 JP 1988192320  
PR 03-AUG-1987 GB 87 8718298, 26-OCT-1987 GB 87 8725052, PR  
13-NOV-1987 GB 87 8726683  
PI NIWA MINEO, SAITO YOSHIMASA, SASAKI HITOSHI, HAYASHI MASAKO,  
PI NOTANI JOJI,  
PI KOBAYASHI MASAKAZU  
PC C12N9/64, C12N1/20, C12N5/00, C12N15/00//A61K37/54, (C12N9/64, PC  
C12R1:19),  
PC (C12N9/64, C12R1:91);  
CC strandedness: Double;  
CC topology: Linear;  
CC hypothetical: No;  
CC \*source: clone=pTTKpadeltatrp;  
FH Key Location/Qualifiers  
FH CDS 1..1068  
FT /product='t-PA'.  
FEATURES  
source Location/Qualifiers  
1..1068  
/organism="synthetic construct"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:32630"  
BASE COUNT 226 a 314 c 313 g 215 t  
ORIGIN  
Query Match 100.0%; Score 1065; DB 6; Length 1068;  
Best Local Similarity 100.0%; Pred. No. 4.7e-248;  
Matches 1065; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 TCTGAGGAAACAGTACTGCTACTTTGGGAATGGGTGAGCTACCGTGGCAACGACAGC 60

Db 4 TCTGAGGGAACAGTGACTGCTTCTTGGGAATGGGTGAGCTACCGTGGGACGACAGC 63  
Qy 61 CTACCGAGTGGGTCCTCTGCTCCCGTGGAAATTCATGATCCTGATAGGCAAGTT 120  
Db 64 CTCACCGAGTGGGTCCTCTGCTCCCGTGGAAATTCATGATCCTGATAGGCAAGTT 123  
Qy 121 TACACAGACAGAACCCAGTGGCCAGGCACTGGGCTGGGCAAAACAATAATTACTGCCGG 180  
Db 124 TACACAGACAGAACCCAGTGGCCAGGCACTGGGCTGGGCAAAACAATAATTACTGCCGG 183  
Qy 181 AATCTGATGGGATGCCAAGCCCTGGTGGCAAGTCTGAGAACCGGAGGCTGACGTGG 240  
Db 184 AATCTGATGGGATGCCAAGCCCTGGTGGCAAGTCTGAGAACCGGAGGCTGACGTGG 243  
Qy 241 GAGTACTGTGATGTCCTCTGCTCCACCTGGCGCCTGAGACACTACAGCCAGCCTCAG 300  
Db 244 GAGTACTGTGATGTCCTCTGCTCCACCTGGCGCCTGAGACACTACAGCCAGCCTCAG 303  
Qy 301 TTTCGCATCAAGAGGAGGCTTTGCGCGACATGCTCCACCCCTGGAGGCTGCCATC 360  
Db 304 TTTCGCATCAAGAGGAGGCTTTGCGCGACATGCTCCACCCCTGGAGGCTGCCATC 363  
Qy 361 TTTCGCAAGCACAGAGGCTGGCCCGGAGAGCGGTTCTGTGGGGGGCATACTATCAGC 420  
Db 364 TTTCGCAAGCACAGAGGCTGGCCCGGAGAGCGGTTCTGTGGGGGGCATACTATCAGC 423  
Qy 421 TCCTGCTGGATTCCTCTGCGGCCCACTGCTTCCAGGAGAGGTTTCGCCCCCAACACCTG 480  
Db 424 TCCTGCTGGATTCCTCTGCGGCCCACTGCTTCCAGGAGAGGTTTCGCCCCCAACACCTG 483  
Qy 481 ACGGTGATCTTTGGGAGAACATAACCGGTGCTCCCTGGCGAGGAGGAGCAAAATTTGAA 540  
Db 484 ACGGTGATCTTTGGGAGAACATAACCGGTGCTCCCTGGCGAGGAGGAGCAAAATTTGAA 543  
Qy 541 CTCGAAATACATGTCATAGGAATTCGATGATGACCTTACGACAATGACATTCG 600  
Db 544 CTCGAAATACATGTCATAGGAATTCGATGATGACCTTACGACAATGACATTCG 603  
Qy 601 CTGCTGACGCTGAATCGGATTCGTCCTGTCGCGAGGAGAGGAGGCTGTCGCACT 660  
Db 604 CTGCTGACGCTGAATCGGATTCGTCCTGTCGCGAGGAGAGGAGGCTGTCGCACT 663  
Qy 661 GTGTGCTTCCCGCGGAGACTGCGAGCTGCGGACTGGAAGGAGGCTCATGTC 720  
Db 664 GTGTGCTTCCCGCGGAGACTGCGAGCTGCGGACTGGAAGGAGGCTCATGTC 723  
Qy 721 TACGGCAGCATGAGGCTTGTCTCTTCTATTCGAGCGGCTGAGGAGGCTCATGTC 780  
Db 724 TACGGCAGCATGAGGCTTGTCTCTTCTATTCGAGCGGCTGAGGAGGCTCATGTC 783  
Qy 781 AGACTGTACCCATCCAGCGCTGCATCAACAATTTACTTTACAGAACAGTCAACGAC 840  
Db 784 AGACTGTACCCATCCAGCGCTGCATCAACAATTTACTTTACAGAACAGTCAACGAC 843  
Qy 841 AACATGCTGTGTGTGAGACACTCGGAGCGCGCGGCCCGGCAAACTTGCACGAGCC 900  
Db 844 AACATGCTGTGTGTGAGACACTCGGAGCGCGCGGCCCGGCAAACTTGCACGAGCC 903  
Qy 901 TCCAGGCGGATTCGGGAGGCGCCCTGTGTGTGTGTAACGATGGCGGATGACTTTGGTG 960  
Db 904 TCCAGGCGGATTCGGGAGGCGCCCTGTGTGTGTGTGTAACGATGGCGGATGACTTTGGTG 963  
Qy 961 GGCAATCATGCTGGGCTGGGCTGTGACAGAGGATGTCGGGCTGTGTACAAAG 1020  
Db 964 GGCAATCATGCTGGGCTGGGCTGTGACAGAGGATGTCGGGCTGTGTACAAAG 1023  
Qy 1021 GTTACCAACTACCTAGATGGAATTCGTGACAACTGCGACCGTGA 1065  
Db 1024 GTTACCAACTACCTAGATGGAATTCGTGACAACTGCGACCGTGA 1068

RESULT 6

A27727 LOCUS A27727 DNA 1170 bp linear PAT 09-JUL-2002  
DEFINITION DNA sequence of coding region in pmTtk.  
ACCESSION A27727  
VERSION A27727.1 GI:21727237  
KEYWORDS synthetic construct  
SOURCE synthetic construct  
ORGANISM artificial construct  
REFERENCE 1 (bases 1 to 1170)  
AUTHORS Niwa, M., Saito, Y., Sasaki, H., Hayashi, M., Notani, J. and Kobayashi, M.  
TITLE New tissue plasminogen activator  
JOURNAL Patent: EP 0302456-A 51 08-FEB-1989;  
FUJISAWA PHARMACEUTICAL CO., LTD  
FEATURES  
Location/Qualifiers  
1..1170  
/organism="synthetic construct"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:32630"  
BASE COUNT 249 a 338 c 346 g 237 t  
ORIGIN  
Query Match 100.0%; Score 1065; DB 6; Length 1170;  
Best Local Similarity 100.0%; Pred. No. 4.8e-248;  
Matches 1065; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 TCTCAGGGAACAGTGACTGCTTCTTGGGAATGGGTGAGCTACCGTGGGACGACAGC 60  
Db 106 TCTCAGGGAACAGTGACTGCTTCTTGGGAATGGGTGAGCTACCGTGGGACGACAGC 165  
Qy 61 CTCACCGAGTGGGTCCTCTGCTCCCGTGGAAATTCATGATCCTGATAGGCAAGTT 120  
Db 166 CTCACCGAGTGGGTCCTCTGCTCCCGTGGAAATTCATGATCCTGATAGGCAAGTT 225  
Qy 121 TACACAGACAGAACCCAGTGGCCAGGCACTGGGCTGGGCAAAACAATAATTACTGCCGG 180  
Db 226 TACACAGACAGAACCCAGTGGCCAGGCACTGGGCTGGGCAAAACAATAATTACTGCCGG 285  
Qy 181 AATCTGATGGGATGCCAAGCCCTGGTGGCAAGTCTGAGAACCGGAGGCTGACGTGG 240  
Db 286 AATCTGATGGGATGCCAAGCCCTGGTGGCAAGTCTGAGAACCGGAGGCTGACGTGG 345  
Qy 241 GAGTACTGTGATGTCCTCTGCTCCACCTGCGGCTGAGACAGTACAGCCAGCCTCAG 300  
Db 346 GAGTACTGTGATGTCCTCTGCTCCACCTGCGGCTGAGACAGTACAGCCAGCCTCAG 405  
Qy 301 TTTCGCATCAAGAGGAGGCTTTGCGCGACATCGCCTCCACCCCTGGCAGGCTGCCATC 360  
Db 406 TTTCGCATCAAGAGGAGGCTTTGCGCGACATCGCCTCCACCCCTGGCAGGCTGCCATC 465  
Qy 361 TTTCGCAAGCACAGAGGCTGGCCCGGAGAGCGGTTCTGTGGGGGGCATACTCATCAGC 420  
Db 466 TTTCGCAAGCACAGAGGCTGGCCCGGAGAGCGGTTCTGTGGGGGGCATACTCATCAGC 525  
Qy 421 TCCTGTGGAATTCCTCTGCGGCCCACTGCTTCCAGGAGAGGTTTCGCCCCCAACACCTG 480  
Db 526 TCCTGTGGAATTCCTCTGCGGCCCACTGCTTCCAGGAGAGGTTTCGCCCCCAACACCTG 585  
Qy 481 ACGGTGATCTTTGGGCAAAACAATAACCGGTCCTGCGGAGGAGGAGCAAAATTTGAA 540  
Db 586 ACGGTGATCTTTGGGCAAAACAATAACCGGTCCTGCGGAGGAGGAGCAAAATTTGAA 645  
Qy 541 GTCGAAAAATACATGTCATAGGAATTCGATGATGACACTTACGACAATGACATTCGCG 600  
Db 646 GTCGAAAAATACATGTCATAGGAATTCGATGATGACACTTACGACAATGACATTCGCG 705  
Qy 601 CTGCTGAGCTGAAATCGGATTCGTCGCCGCTGTGCCCGAGGAGCAGCGTGGTCCGCACT 660  
Db 706 CTGCTGAGCTGAAATCGGATTCGTCGCCGCTGTGCCCGAGGAGCAGCGTGGTCCGCACT 765  
Qy 661 GTGTGCTTCCCGCGGAGACTGCGAGCTGCCGAGCTGGACGGAGTGTGAGCTCTCCGGC 720



```
Db 1126 GTTACCAACTACCTAGATGGATTCTGACACATGCGACCGTGA 1170
|||||
A27435 1314 bp DNA linear PAT 09-JUL-2002
DEFINITION DNA sequence of coding region in pTQkPAdeltatrp.
ACCESSION A27435
VERSION A27435.1 GI:21727230
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
REFERENCE 1 (bases 1 to 1314)
AUTHORS Niwa,M., Saito,Y., Sasaki,H., Hayashi,M., Notani,J. and
Kobayashi,M.
TITLE New tissue plasminogen activator
JOURNAL Patent: EP 0302456-A 35 08-FEB-1989;
FUJISAWA PHARMACEUTICAL CO., LTD
FEATURES
Location/Qualifiers
1..1314
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
BASE COUNT 286 a 386 c 393 g 249 t
ORIGIN
Query Match 100.0%; Score 1065; DB 6; Length 1314;
Best Local Similarity 100.0%; Pred. No. 4.8e-248;
Matches 1065; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TCTGAGGAAACAGTGACTGCTACTTTGGGAATGGGTGAGCTACCGTGGCAGCGACACAGC 60
Db 250 TCTGAGGAAACAGTGACTGCTACTTTGGGAATGGGTGAGCTACCGTGGCAGCGACACAGC 309
QY 61 CTCACCGAGTGGGTGCTCTCTGCTCCGCTGGGAATTCATGATCCTGATAGGCAAGTT 120
Db 310 CTCACCGAGTGGGTGCTCTCTGCTCCGCTGGGAATTCATGATCCTGATAGGCAAGTT 369
QY 121 TACACAGACAGAACCCAGTGCACGACCTGGGCTGGGCAACATAATCTACGCGG 180
Db 370 TACACAGACAGAACCCAGTGCACGACCTGGGCTGGGCAACATAATCTACGCGG 429
QY 181 AATCCTGATGGGATGCCAAGCCCTGTGTCACAGTGTGTAAGAACCGCAGGCTGACGTGG 240
Db 430 AATCCTGATGGGATGCCAAGCCCTGTGTCACAGTGTGTAAGAACCGCAGGCTGACGTGG 489
QY 241 GAGTACTGTGATGTCCTCTCTGCTCCACCTGGGCTGGGCAACATAATCTACGCGG 300
Db 490 GAGTACTGTGATGTCCTCTCTGCTCCACCTGGGCTGGGCAACATAATCTACGCGG 549
QY 301 TTTCGCATCAAGAGAGGCTCTTCCGCGACATCGCTCCACCCCTGGCAGGCTGCCATC 360
Db 550 TTTCGCATCAAGAGAGGCTCTTCCGCGACATCGCTCCACCCCTGGCAGGCTGCCATC 609
QY 361 TTTCGCATCAAGAGAGGCTCTTCCGCGACATCGCTCCACCCCTGGCAGGCTGCCATC 420
Db 610 TTTCGCATCAAGAGAGGCTCTTCCGCGACATCGCTCCACCCCTGGCAGGCTGCCATC 669
QY 421 TCCTGCTGGATTCCTCTGTCGCGCCCACTGCTTCCAGGAGAGGTTTCCGCGCCCAACCTG 480
Db 670 TCCTGCTGGATTCCTCTGTCGCGCCCACTGCTTCCAGGAGAGGTTTCCGCGCCCAACCTG 729
QY 481 ACGGTGATCTTGGGAGAGAGGCTCTTCCGCGACATCGCTCCGCGAGGAGGAGGAGGAGG 540
Db 730 ACGGTGATCTTGGGAGAGAGGCTCTTCCGCGACATCGCTCCGCGAGGAGGAGGAGGAGG 789
QY 541 GTCGAAAATACATTTGTCATTAAGGAATTCGATGATGACCTTACGCAATGACATTCGCG 600
Db 790 GTCGAAAATACATTTGTCATTAAGGAATTCGATGATGACCTTACGCAATGACATTCGCG 849
QY 601 CTGTGCGAGCTGAATCGGATTCGTCGCCGCTGTGTCGCCAGGAGAGCAGCGTGGTCCGCACT 660
```

```
Db 850 CTGCTGCAGCTGAATCGGATTCGTCGCGCTGTGCCAGAGAGCAGCGTGGTCCGCACT 909
|||||
QY 661 GTGTGCCCTTCCCGCGGACCTGCGAGCTGCGGACCTGCGAGCGAGTGTGAGCTCTCCGCG 720
|||||
Db 910 GTGTGCCCTTCCCGCGGACCTGCGAGCTGCGGACCTGCGAGCGAGTGTGAGCTCTCCGCG 969
|||||
QY 721 TACGGCAAGCATAGGCGCTTCTCTCTTTTATTTCGGAGCGGCTGAAGAGGCTCATGTC 780
Db 970 TACGGCAAGCATAGGCGCTTCTCTCTTTTATTTCGGAGCGGCTGAAGAGGCTCATGTC 1029
|||||
QY 781 AGACTGTACCCATCCAGCGGCTGCACATCAACAATTTACTTTACAGAACAGTCAACGAC 840
|||||
Db 1030 AGACTGTACCCATCCAGCGGCTGCACATCAACAATTTACTTTACAGAACAGTCAACGAC 1089
|||||
QY 841 AACATGCTGTGCTGGAGACACTCGGAGCGGCGGCCAGGCAAACTTGCACGACGCC 900
|||||
Db 1090 AACATGCTGTGCTGGAGACACTCGGAGCGGCGGCCAGGCAAACTTGCACGACGCC 1149
|||||
QY 901 TGCAGGCGCATTCGGAGGCGCCCTGCTGTGTCTGAACGATGGCGCGCATGACTTTGGTG 960
Db 1150 TGCAGGCGCATTCGGAGGCGCCCTGCTGTGTCTGAACGATGGCGCGCATGACTTTGGTG 1209
|||||
QY 961 GGCATCATCAGCTGGGCGCTGGGCTGTGGACAGAGGATGTCCCGGCTGTGTACACAAAG 1020
Db 1210 GGCATCATCAGCTGGGCGCTGGGCTGTGGACAGAGGATGTCCCGGCTGTGTACACAAAG 1269
|||||
QY 1021 GTTACCAACTACCTAGACTGGATTCTGTCACAACTGCGACCGTGA 1065
Db 1270 GTTACCAACTACCTAGACTGGATTCTGTCACAACTGCGACCGTGA 1314
|||||
RESULT 9
E01937
LOCUS E01937 1314 bp DNA linear PAT 29-SEP-1997
DEFINITION Synthetic DNA encoding new t-PA tissue plasminogen activator.
ACCESSION E01937
VERSION E01937.1 GI:2170185
KEYWORDS JP 1989104167-A/4.
SOURCE synthetic construct
ORGANISM synthetic construct
REFERENCE 1 (bases 1 to 1314)
AUTHORS Niwa,M., Saito,Y., Sasaki,H., Hayashi,M., Notani,J. and
Kobayashi,M.
TITLE NOVEL TISSUE PLASMINOGEN ACTIVATION FACTOR
JOURNAL Patent: JP 1989104167-A 4 21-APR-1989;
FUJISAWA PHARMACEUT CO LTD
COMMENT OS Artificial gene
OC Artificial sequence; Genes.
PN JP 1989104167-A/4
PD 21-APR-1989
PF 01-AUG-1988 JP 1988192320
PR 03-AUG-1987 GB 87 8718298, 26-OCT-1987 GB 87 8725052, PR
13-NOV-1987 GB 87 8726683
PI NIWA MINEO, SAITO YOSHIMASA, SASAKI HITOSHI, HAYASHI MASAKO,
PI NOTANI JOJI,
PI KOBAYASHI MASAKAZU
PC C12N9/64, C12N1/20, C12N5/00, C12N15/00//A61K37/54, (C12N9/64, PC
C12R1:19),
PC (C12N9/64, C12R1:91);
CC strandedness: Double;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
CC *source: clone=pTQkPAdeltatrp;
FH Key Location/Qualifiers
FT CDS 1..1314
FT /product='t-PA'.
FT Location/Qualifiers
1..1314
FT /organism="synthetic construct"
```

```
/mol_type="genomic DNA"
/db_xref="taxon:32630"
BASE COUNT 286 a 386 c 393 g 249 t
ORIGIN
1 TCTGAGGAAACAGTACTGCTACTTTGGGAATGGGTCAAGCTACCGTGGCAACGACAGC 60
250 TCTGAGGAAACAGTACTGCTACTTTGGGAATGGGTCAAGCTACCGTGGCAACGACAGC 309
61 CTCACCGAGTGGGCTCTCTGCTCCCTCCCTGGGAATTCATGATCCTCATAGGCAAGTT 120
310 CTCACCGAGTGGGCTCTCTGCTCCCTCCCTGGGAATTCATGATCCTCATAGGCAAGTT 369
121 TACACAGACAGAAACCCAGTGGCCAGGACCTGGGCTGGGCAACATAATTAATCTGCGG 180
370 TACACAGACAGAAACCCAGTGGCCAGGACCTGGGCTGGGCAACATAATTAATCTGCGG 429
181 AATCTGATGGGATGCCAGCCCTGGTGGCCAGCTGTGAAGAACCGCAGCTGACGTGG 240
430 AATCTGATGGGATGCCAGCCCTGGTGGCCAGCTGTGAAGAACCGCAGCTGACGTGG 489
241 GAGTACTGTGATGTCCTCTCTGCTCCACCTGGCCCTGAGACAGTACAGCCAGCTCAG 300
490 GAGTACTGTGATGTCCTCTCTGCTCCACCTGGCCCTGAGACAGTACAGCCAGCTCAG 549
301 TTTCGCATCAAGAGGGGCTTTTCGCCGACATCGCTCCACCCCTGGCAGGCTGCCATC 360
550 TTTCGCATCAAGAGGGGCTTTTCGCCGACATCGCTCCACCCCTGGCAGGCTGCCATC 609
361 TTTCGCATCAAGAGGGTGGCCGAGAGGGTTCCTGTGGGGGGCATACTATCAGC 420
610 TTTCGCATCAAGAGGGTGGCCGAGAGGGTTCCTGTGGGGGGCATACTATCAGC 669
421 TCCTGCTGATTTCTCTGTGGCCCACTGCTTCCAGGAGAGGTTTCGCCCCACCCACTG 480
670 TCCTGCTGATTTCTCTGTGGCCCACTGCTTCCAGGAGAGGTTTCGCCCCACCCACTG 729
481 ACGGTGATCTTTGGGAGAACATACCGGGTGGTCCCTGGCGAGGAGGAGAGAAATTTGAA 540
730 ACGGTGATCTTTGGGAGAACATACCGGGTGGTCCCTGGCGAGGAGGAGAGAAATTTGAA 789
541 GTCGAAATATACATTTGTCATAGGAATTCGATGATGACATTCAGCAATGACATTCGG 600
790 GTCGAAATATACATTTGTCATAGGAATTCGATGATGACATTCAGCAATGACATTCGG 849
601 CTGCTGACGTGAATTCGATTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
850 CTGCTGACGTGAATTCGATTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 909
661 GTGTCCTTCCCGCGGACCTGAGCTGCGGACTGCGAGAGTGTGAGCTCTCGGC 720
910 GTGTCCTTCCCGCGGACCTGAGCTGCGGACTGCGAGAGTGTGAGCTCTCGGC 969
721 TACGCAAGCATGAGGCTTGTCTCTTTCTATTCGAGCGGCTGAGAGAGCTCATGTC 780
970 TACGCAAGCATGAGGCTTGTCTCTTTCTATTCGAGCGGCTGAGAGAGCTCATGTC 1029
781 AGACTGTACCCATCCAGCGCTGACATCACAACTTTTAAACAGAACAGTCAACGAC 840
1030 AGACTGTACCCATCCAGCGCTGACATCACAACTTTTAAACAGAACAGTCAACGAC 1089
841 AACATGCTGTGTGAGACATCCGAGCGCGCGCCCGAGGCAAACTTGACAGAGCC 900
1090 AACATGCTGTGTGAGACATCCGAGCGCGCGCGCCCGAGGCAAACTTGACAGAGCC 1149
901 TGCCAGGGGATTCGGGAGGCGCCCTGCTGTGCTGAACGATGGCGCATGCTTTGGTG 960
1150 TGCCAGGGGATTCGGGAGGCGCCCTGCTGTGCTGAACGATGGCGCATGCTTTGGTG 1209
```

---

```
Qy 961 GGCAATCATCAGCTGGGCTGGGCTGTGTGGAAGAGATGTCCGGGTGTGTACACAAAG 1020
Db 1210 GGCAATCATCAGCTGGGCTGGGCTGTGTGGAAGAGATGTCCGGGTGTGTACACAAAG 1269
Qy 1021 GTTACCAACTACCTAGACTGGATTCGTGACAAACATGCGACCGTGA 1065
Db 1270 GTTACCAACTACCTAGACTGGATTCGTGACAAACATGCGACCGTGA 1314

RESULT 10
A27725
LOCUS DNA sequence of coding region in pmtQk112. 1419 bp DNA linear PAT 09-JUL-2002
ACCESSION A27725
VERSION A27725.1 GI:21727236
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM artificial sequences.
REFERENCE 1 (bases 1 to 1419)
AUTHORS Niwa,M., Saito,Y., Sasaki,H., Hayashi,M., Notani,J. and Kobayashi,M.
TITLE New tissue plasminogen activator
JOURNAL Patent: EP 0302456-A 49 08-FEB-1989; FUJISAWA PHARMACEUTICAL CO., LTD
FEATURES
source
1..1419
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
BASE COUNT 309 a 413 c 426 g 271 t
ORIGIN

Query Match 100.0%; Score 1065; DB 6; Length 1419;
Best Local Similarity 100.0%; Pred. No. 4.8e-248;
Matches 1065; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 TCTGAGGAAACAGTACTGCTACTTTGGGAATGGGTCAAGCTACCGTGGCAACGACAGC 60
Db 355 TCTGAGGAAACAGTACTGCTACTTTGGGAATGGGTCAAGCTACCGTGGCAACGACAGC 414
Qy 61 CTCACCGAGTGGGTGGCTCTCTGCTCCCTGGTGAATTCATGATCCTGATAGGCAAGTT 120
Db 415 CTCACCGAGTGGGTGGCTCTCTGCTCCCTGGTGAATTCATGATCCTGATAGGCAAGTT 474
Qy 121 TACACAGACAGAAACCCAGTGGCCAGGACCTGGGCTGGGCAACATAATTAATCTGCGG 180
Db 475 TACACAGACAGAAACCCAGTGGCCAGGACCTGGGCTGGGCAACATAATTAATCTGCGG 534
Qy 181 AATCTGATGGGATGCCAGCCCTGGTGGCCAGCTGTGAAGAACCGCAGCTGACGTGG 240
Db 535 AATCTGATGGGATGCCAGCCCTGGTGGCCAGCTGTGAAGAACCGCAGCTGACGTGG 594
Qy 241 GAGTACTGTGATGTCCTCTCTGCTCCACCTGGCGCTGAGACAGTACAGCCAGCTCAG 300
Db 595 GAGTACTGTGATGTCCTCTCTGCTCCACCTGGCGCTGAGACAGTACAGCCAGCTCAG 654
Qy 301 TTTCGCATCAAGAGGGCTTTTCGCCGACATCGCTCCACCCCTGGCAGGCTGCCATC 360
Db 655 TTTCGCATCAAGAGGGCTTTTCGCCGACATCGCTCCACCCCTGGCAGGCTGCCATC 714
Qy 361 TTTCGCATCAAGAGGGCTTTTCGCCGACATCGCTCCACCCCTGGCAGGCTGCCATC 420
Db 715 TTTCGCATCAAGAGGGCTTTTCGCCGACATCGCTCCACCCCTGGCAGGCTGCCATC 774
Qy 421 TCCTGCTGATTTCTCTGTGGCGCCACTGTCTTCAGGAGAGGTTTCGCCCCACCACTG 480
Db 775 TCCTGCTGATTTCTCTGTGGCGCCACTGTCTTCAGGAGAGGTTTCGCCCCACCACTG 834
Qy 481 ACGGTGATCTTTGGGAGAACATACCGGGTGGTCCCTGGCGAGGAGGAGAAATTTGAA 540
Db 835 ACGGTGATCTTTGGGAGAACATACCGGGTGGTCCCTGGCGAGGAGGAGAAATTTGAA 894
```

QY 541 GTCGAAATATACATTGTCCATAGGAATTCGATGATGACATTCACGACAATGACATTGGC 600  
 Db 895 GTCGAAATATACATTGTCCATAGGAATTCGATGATGACATTCACGACAATGACATTGGC 954  
 QY 601 CTGCTGCAGCTGAATCGAATTCGTCCTGTCGCCAGGAGAGCAGGTCGTCGCACT 660  
 Db 955 CTGCTGCAGCTGAATCGAATTCGTCCTGTCGCCAGGAGAGCAGGTCGTCGCACT 1014  
 QY 661 GTGTGCTTCCCGCGGAGACTGAGCTGCGGACTGCGAGCTGCGAGCTGTGAGCTCTCCGGC 720  
 Db 1015 GTGTGCTTCCCGCGGAGACTGAGCTGCGGACTGCGAGCTGTGAGCTCTCCGGC 1074  
 QY 721 TACGCAAGCATAGGCTTGTCTCTTTCTATTCGAGCGGCTGAAGAGGCTCATGTC 780  
 Db 1075 TACGCAAGCATAGGCTTGTCTCTTTCTATTCGAGCGGCTGAAGAGGCTCATGTC 1134  
 QY 781 AGACTGTACCCATCCAGCGCTGCACATCAACAATTTACTTTAAACAGAACAGTCAACGAC 840  
 Db 1135 AGACTGTACCCATCCAGCGCTGCACATCAACAATTTACTTTAAACAGAACAGTCAACGAC 1194  
 QY 841 AACATGCTGTGTGCTGGAGACATCGAGCGCGCGGCCCGGCAAACTTGCACGACGCC 900  
 Db 1195 AACATGCTGTGTGCTGGAGACATCGAGCGCGCGGCCCGGCAAACTTGCACGACGCC 1254  
 QY 901 TCCAGGCGGATTCGGAGGCGCCCTGTCGTCGTAACGATGCGCGGATGCTTTGGTG 960  
 Db 1255 TCCAGGCGGATTCGGAGGCGCCCTGTCGTCGTAACGATGCGCGGATGCTTTGGTG 1314  
 QY 961 GGCATCATCAGTGGGCTCGGAGGCGCCCTGTCGTCGTAACGATGCGCGGATGCTTTGGTG 1020  
 Db 1315 GGCATCATCAGTGGGCTCGGAGGCGCCCTGTCGTCGTAACGATGCGCGGATGCTTTGGTG 1374  
 QY 1021 GTTACCACTACCTAGACTGGATTTCGTGACCAACATCGCACCGTGA 1065  
 Db 1375 GTTACCACTACCTAGACTGGATTTCGTGACCAACATCGCACCGTGA 1419

RESULT 11  
 AR059996  
 LOCUS AR059996 1419 bp DNA linear PAT 29-SEP-1999  
 DEFINITION Sequence 62 from patent US 5840533.  
 ACCESSION AR059996  
 VERSION AR059996.1 GI:5986446  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unknown.  
 REFERENCE  
 AUTHORS Niwa, M., Saico, Y., Sasaki, H., Hayashi, M., Notani, J. and Kobayashi, M.  
 TITLE Tissue plasminogen activator  
 JOURNAL Patent: US 5840533-A 62 24-NOV-1998;  
 FEATURES  
 source 1..1419  
 BASE COUNT 309 a 416 c 423 g 271 t  
 ORIGIN

Query Match 100.0%; Score 1065; DB 6; Length 1419;  
 Best Local Similarity 100.0%; Pred. No. 4.8e-248;  
 Matches 1065; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCTGAGGAAACAGTACTGCTACTTTGGGAATGGGTGAGCTACCGTGGCAGCACGACAGC 60  
 Db 355 TCTGAGGAAACAGTACTGCTACTTTGGGAATGGGTGAGCTACCGTGGCAGCACGACAGC 414  
 QY 61 CTCACGAGTCGGTGCTCTGCTCCCTGCGAATTCATGATCTCTGATAGGCAAGTT 120  
 Db 415 CTCACGAGTCGGTGCTCTGCTCCCTGCGAATTCATGATCTCTGATAGGCAAGTT 474  
 QY 121 TACACGACAGAACCCAGTCCCGAGGCTGCGGCTGGGCAACATAATTAATCTGCGG 180  
 Db 475 TACACGACAGAACCCAGTCCCGAGGCTGCGGCTGGGCAACATAATTAATCTGCGG 534

QY 181 AATCTGATGGGATGCCAAGCCCTGCTGTCACCGTGTGAAGAAACCGCAGGCTGACGTGG 240  
 Db 535 AATCTGATGGGATGCCAAGCCCTGCTGTCACCGTGTGAAGAAACCGCAGGCTGACGTGG 594  
 QY 241 GAGTACTGTGATGTCCTCTCTGCTCCACCTGCGGCTGAGACAGTACAGCCAGGCTCAG 300  
 Db 595 GAGTACTGTGATGTCCTCTCTGCTCCACCTGCGGCTGAGACAGTACAGCCAGGCTCAG 654  
 QY 301 TTTCCGATCAAAAGAGGAGGCTCTTCGCGGACATCGCTCCCAACCTGCGGCTGCGCATC 360  
 Db 655 TTTCCGATCAAAAGAGGAGGCTCTTCGCGGACATCGCTCCCAACCTGCGGCTGCGCATC 714  
 QY 361 TTTCCGATCAAAAGAGGAGGCTCTTCGCGGACATCGCTCCCAACCTGCGGCTGCGCATC 420  
 Db 715 TTTCCGATCAAAAGAGGAGGCTCTTCGCGGACATCGCTCCCAACCTGCGGCTGCGCATC 774  
 QY 421 TCTGCTGGATTCCTCTGCGGCGGCTCTTCGAGGAGGCTTTCCGCGGCGGCTGCGCATC 480  
 Db 775 TCTGCTGGATTCCTCTGCGGCGGCTCTTCGAGGAGGCTTTCCGCGGCGGCTGCGCATC 834  
 QY 481 ACAGTGTATCTTGGGACAGAAATACCGGCTGCTCCCTGCGGAGGAGGAGCAAAATTTCAA 540  
 Db 835 ACAGTGTATCTTGGGACAGAAATACCGGCTGCTCCCTGCGGAGGAGGAGCAAAATTTCAA 894  
 QY 541 GTCGAAAATACATTTGTCATTAAGGAATTCGATGATGACACTTACGACAATGACATTGCG 600  
 Db 895 GTCGAAAATACATTTGTCATTAAGGAATTCGATGATGACACTTACGACAATGACATTGCG 954  
 QY 601 CTGCTGAGCTGAATTCGATTCCTGCTGTCGCCAGGAGAGGCTGCTGCGGCTGCGCATC 660  
 Db 955 CTGCTGAGCTGAATTCGATTCCTGCTGTCGCCAGGAGAGGCTGCTGCGGCTGCGCATC 1014  
 QY 661 GTGTGCTTCCCGCGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720  
 Db 1015 GTGTGCTTCCCGCGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1074  
 QY 721 TACGCAAGCATAGGCTTGTCTCTTTCTATTCGAGCGGCTGAAGAGGCTCATGTC 780  
 Db 1075 TACGCAAGCATAGGCTTGTCTCTTTCTATTCGAGCGGCTGAAGAGGCTCATGTC 1134  
 QY 781 AGACTGTACCCATCCAGCGCTGCACATCAACAATTTACTTTAAACAGAACAGTCAACGAC 840  
 Db 1135 AGACTGTACCCATCCAGCGCTGCACATCAACAATTTACTTTAAACAGAACAGTCAACGAC 1194  
 QY 841 AACATGCTGTGTGCTGGAGACATCGAGCGCGCGGCCCGGCAAACTTGCACGACGCC 900  
 Db 1195 AACATGCTGTGTGCTGGAGACATCGAGCGCGCGGCCCGGCAAACTTGCACGACGCC 1254  
 QY 901 TCCAGGCGGATTCGGAGGCGCCCTGTCGTCGTAACGATGCGCGGATGCTTTGGTG 960  
 Db 1255 TCCAGGCGGATTCGGAGGCGCCCTGTCGTCGTAACGATGCGCGGATGCTTTGGTG 1314  
 QY 961 GGCATCATCAGTGGGCTCGGAGGCGCCCTGTCGTCGTAACGATGCGCGGATGCTTTGGTG 1020  
 Db 1315 GGCATCATCAGTGGGCTCGGAGGCGCCCTGTCGTCGTAACGATGCGCGGATGCTTTGGTG 1374  
 QY 1021 GTTACCACTACCTAGACTGGATTTCGTGACCAACATCGCACCGTGA 1065  
 Db 1375 GTTACCACTACCTAGACTGGATTTCGTGACCAACATCGCACCGTGA 1419

RESULT 12  
 E01944  
 LOCUS E01944  
 DEFINITION Synthetic DNA encoding new t-PA tissue plasminogen activator.  
 ACCESSION E01944  
 VERSION E01944.1 GI:2170192  
 KEYWORDS JP 1989104167-A/11.  
 SOURCE synthetic construct  
 ORGANISM artificial sequences.  
 REFERENCE 1 (bases 1 to 1419)







Db	745	TACACAGACAGAACCCAGGTGCCAGGCACTGGCGCTTGGGCAACATAAATTACTGCGGG	804
Qy	181	AATCCTGATGGGGATGCCAAGCCCTTGGTCCACAGTGTCTGAAGAACCGCAGGCTACGCTGG	240
Db	805	AATCCTGATGGGGATGCCAAGCCCTTGGTCCACAGTGTCTGAAGAACCGCAGGCTACGCTGG	864
Qy	241	GAGTACTGTGATGTGCCCTCTGTCTCCACTGTCGGGCTTGACACAGTACAGGCAGGCTCAG	300
Db	865	GAGTACTGTGATGTGCCCTCTGTCTCCACTGTCGGGCTTGACACAGTACAGGCAGGCTCAG	924
Qy	301	TTTTCGCATCAAAAGAGGGCTCTTCGCGGACATCGCCTCCACCCCTGCGAGGCTGCCATC	360
Db	925	TTTTCGCATCAAGAGGGCTCTTCGCGGACATCGCCTCCACCCCTGGCAGGCTGCCATC	984
Qy	361	TTTTCGCAAGCACAGGAGGTTCGCCGAGAGCGGTTCTGTGTCGGGGGACATACTCATCAGC	420
Db	985	TTTTCGCAAGCACAGGAGGTTCGCCGAGAGCGGTTCTGTGTCGGGGGACATACTCATCAGC	1044
Qy	421	TCTGCTGGGATTCCTCTGTCGGCCCACTGCTTCCAGGAGAGGTTTTCGCCCCACCACTG	480
Db	1045	TCTGCTGGGATTCCTCTGTCGGCCCACTGCTTCCAGGAGAGGTTTTCGCCCCACCACTG	1104
Qy	481	ACGGTGAATCTTTGGGCAGAACATACCGGGTGTCTCCGCGAGGAGGACAGAAATTTGAA	540
Db	1105	ACGGTGAATCTTTGGGCAGAACATACCGGGTGTCTCCGCGAGGAGGACAGAAATTTGAA	1164
Qy	541	GTGCAAAATATCATTTGTCATAGGAATTCGATGATGACATTAACGACAATGACATTCGC	600
Db	1165	GTGCAAAATATCATTTGTCATAGGAATTCGATGATGACATTAACGACAATGACATTCGC	1224
Qy	601	CTGCTGCAGCTGAATCGGATTCGTCGCCGTGTCGCCAGAGAGCAGCGTGGTCCGCACCT	660
Db	1225	CTGCTGCAGCTGAATCGGATTCGTCGCCGTGTCGCCAGAGAGCAGCGTGGTCCGCACCT	1284
Qy	661	GTGTGCTTCCCGCGCGACCTGACGTGCGGACTGCGAGTGTGAGCTCTCCGGC	720
Db	1285	GTGTGCTTCCCGCGCGACCTGACGTGCGGACTGCGAGTGTGAGCTCTCCGGC	1344
Qy	721	TACGGCAAGCATGAGGCTTGTCTCTTTCTATTTCGGAGCGGCTGAAGGAGGCTCATGTC	780
Db	1345	TACGGCAAGCATGAGGCTTGTCTCTTTCTATTTCGGAGCGGCTGAAGGAGGCTCATGTC	1404
Qy	781	AGACTGTACCCTATCCAGCCGCTGCACATCAACATTTACTTTAACAGAACAGTCACCGAC	840
Db	1405	AGACTGTACCCTATCCAGCCGCTGCACATCAACATTTACTTTAACAGAACAGTCACCGAC	1464
Qy	841	AACATGCTGTGTGTGGAGACATTCGAGCGGGGGCCCCAGGCAAACTTTCACAGCGCC	900
Db	1465	AACATGCTGTGTGTGGAGACATTCGAGCGGGGGCCCCAGGCAAACTTTCACAGCGCC	1524
Qy	901	TGCCAGGGGATTTGGGAGGCGCCCTGGTGTGTGTAACGATGGCCGATGACTTTGGTG	960
Db	1525	TGCCAGGGGATTTGGGAGGCGCCCTGGTGTGTGTAACGATGGCCGATGACTTTGGTG	1584
Qy	961	GGCATCATCAGCTGGGCGCTGGGCTGTGGGACAGAAAGGATGTCGCCGGTGTGTACACAAG	1020
Db	1585	GGCATCATCAGCTGGGCGCTGGGCTGTGGGACAGAAAGGATGTCGCCGGTGTGTACACAAG	1644
Qy	1021	GTTTACCAATCTACCTAGACTGGATTTGTGTGACAAATGCGACCGTGA	1065
Db	1645	GTTTACCAATCTACCTAGACTGGATTTGTGTGACAAATGCGACCGTGA	1689

RESULT	I06614
LOCUS	I06614
DEFINITION	I06614 Sequence 46 from Patent EP 0293934.
ACCESSION	I06614
VERSION	I06614.1 GI:590640
KEYWORDS	.
SOURCE	Unknown.
ORGANISM	Unknown.
DNA	linear
PAT	02-DEC-1994

```
Db 1465 AACATGCTGTGTCTGGAGACACTCGGAGCGCGGCCCCAGGCAAACTTGACGACGCC 1524
Qy 901 TCCAGGCGGATTCGGGAGGCCCTCTGTGTCTGAACGATGCGCGCATGACTTTGGTG 960
Db 1525 TCCAGGCGGATTCGGGAGGCCCTCTGTGTCTGAACGATGCGCGCATGACTTTGGTG 1584
Qy 961 GGCATCATCAGCTGGGGCTGGGCTGTGGACAGAGGATGTCCCGGGTGTGTACACAAG 1020
Db 1585 GGCATCATCAGCTGGGGCTGGGCTGTGGACAGAGGATGTCCCGGGTGTGTACACAAG 1644
Qy 1021 GTTACCAACTACTAGACTGGATTGTGACAAACATGCGACCGTGA 1065
Db 1645 GTTACCAACTACTAGACTGGATTGTGACAAACATGCGACCGTGA 1689

RESULT 15
108789
LOCUS 108789 1689 bp DNA linear PAT 02-DEC-1994
DEFINITION Sequence 3 from Patent WO 8804690.
ACCESSION I08789
VERSION I08789.1 GI:588494
KEYWORDS
SOURCE Unknown.
ORGANISM
Unclassified.
REFERENCE
1 (bases 1 to 1689)
AUTHORS Bollen,A.J., Gheysen,D., Jacobs,P., Pierard,L. and Collen,D.J.
JOURNAL Patent: WO 8804690-A 3 30-JUN-1988;
FEATURES
Location/Qualifiers
source 1..1689
/organism="unknown"
BASE COUNT 379 a 480 c 505 g 325 t
ORIGIN

Query Match 100.0%; Score 1065; DB 6; Length 1689;
Best Local Similarity 100.0%; Pred. No. 4.8e-248;
Matches 1065; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCTGAGGAAACAGTACTGTCTACTTTGGGAATGGGTGAGCTACCGCTACCGTGGGACGACAGC 60
Db 625 TCTGAGGAAACAGTACTGTCTACTTTGGGAATGGGTGAGCTACCGTGGGACGACAGC 684
Qy 61 CTCACCGAGTCGGGTGCTCTCTGCTCCGCTGGAAATCCATGATCCTCATAGGCAAGTT 120
Db 685 CTCACCGAGTCGGGTGCTCTCTGCTCCGCTGGAAATCCATGATCCTCATAGGCAAGTT 744
Qy 121 TACACAGCACAGAACCCGAGTCCCGAGGCACTGGGCTTGGGCAAAACATAATTAATCTGCGG 180
Db 745 TACACAGCACAGAACCCGAGTCCCGAGGCACTGGGCTTGGGCAAAACATAATTAATCTGCGG 804
Qy 181 AATCCTGATGGGATGCCAAGCCCTGGTCCACGCTGTGAAGAACCGGAGCTGACGTTG 240
Db 805 AATCCTGATGGGATGCCAAGCCCTGGTCCACGCTGTGAAGAACCGGAGCTGACGTTG 864
Qy 241 GAGTACTGTGATGTCCTCTCTGCTCCACCTGCGGCTGAGACAGTACAGCCAGCTCAG 300
Db 865 GAGTACTGTGATGTCCTCTCTGCTCCACCTGCGGCTGAGACAGTACAGCCAGCTCAG 924
Qy 301 TTTGCGCATCAAGAGGGGCTCTTGCGCGACATCGCTCCACCCCTGGCAGGCTGCCATC 360
Db 925 TTTGCGCATCAAGAGGGGCTCTTGCGCGACATCGCTCCACCCCTGGCAGGCTGCCATC 984
Qy 361 TTTGCGCAGCAGGAGTTCGCGGAGAGCGGTTCTGTGGGGGGGATCTACTCATCAGC 420
Db 985 TTTGCGCAGCAGGAGTTCGCGGAGAGCGGTTCTGTGGGGGGGATCTACTCATCAGC 1044
Qy 421 TCCTGCTGATTTCTCTGTCGCGCCCACTGCTTCCAGAGAGAGGTTTCCGCCCCACCACTG 480
Db 1045 TCCTGCTGATTTCTCTGTCGCGCCCACTGCTTCCAGAGAGAGGTTTCCGCCCCACCACTG 1104
Qy 481 ACGGTGATCTTTGGCAGAACATACCGGGTGTCTCTGCGGAGGAGGACAGAAATTTGAA 540
```

Search completed: August 19, 2003, 04:46:36  
Job time : 4472.78 secs

```
Db 1105 ACGTGTATCTTTGGCAGAACATACCGGGTGTCTCTGCGAGGAGGAGCAGAAATTTGAA 1164
Qy 541 GTCGAAAATATACATTGTTCATTAAGGAATTCGATGATGACACTTACGACAAATGACATTGCG 600
Db 1165 GTCGAAAATATACATTGTTCATTAAGGAATTCGATGATGACACTTACGACAAATGACATTGCG 1224
Qy 601 CTGCTGACAGCTGAAATCGGATTCGTCGCGCTGTGCCAGGAGAGCAGCGTGGTCCGCACT 660
Db 1225 CTGCTGACAGCTGAAATCGGATTCGTCGCGCTGTGCCAGGAGAGCAGCGTGGTCCGCACT 1284
Qy 661 GTGTGCTTCCCGCGGCGACCTGTCAGTCCGAGCTGCGACGAGAGTGTGAGCTCTCCGCGC 720
Db 1285 GTGTGCTTCCCGCGGCGACCTGTCAGTCCGAGCTGCGACGAGTGTGAGCTCTCCGCGC 1344
Qy 721 TAGCGCAAGCATGAGGCTTGTCTCTCTTCTATTTCGAGCGGCTGAAGGAGGCTCATGTC 780
Db 1345 TAGCGCAAGCATGAGGCTTGTCTCTCTTCTATTTCGAGCGGCTGAAGGAGGCTCATGTC 1404
Qy 781 AGACTGTACCCATCCAGCCGCTGCACATCAACAATTTACTTTAAACAGAAACAGTCAACCGAC 840
Db 1405 AGACTGTACCCATCCAGCCGCTGCACATCAACAATTTACTTTAAACAGAAACAGTCAACCGAC 1464
Qy 841 AACATGCTGTGTCTGTGAGACACTCGGAGCGGCGGCGCCAGGCAAACTTGCACGACGCC 900
Db 1465 AACATGCTGTGTCTGTGAGACACTCGGAGCGGCGGCGCCAGGCAAACTTGCACGACGCC 1524
Qy 901 TGCAGGCGGATTCGGGAGGCCCTGCTGCTGCTGAACGATGCGCGCATGACTTTGGTG 960
Db 1525 TGCAGGCGGATTCGGGAGGCCCTGCTGCTGCTGAACGATGCGCGCATGACTTTGGTG 1584
Qy 961 GGCATCATCAGCTGGGCTGCTGGGCTGTGGACAGAGGATGTCCCGGGTGTGTACACAAAG 1020
Db 1585 GGCATCATCAGCTGGGCTGCTGGGCTGTGGACAGAGGATGTCCCGGGTGTGTACACAAAG 1644
Qy 1021 GTTACCAACTACTCTAGACTGGATTCTGTAACAACATGCGACCGTGA 1065
Db 1645 GTTACCAACTACTCTAGACTGGATTCTGTAACAACATGCGACCGTGA 1689
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 18, 2003, 19:51:53 ; Search time 375.192 Seconds  
(without alignments)  
7662.485 Million cell updates/sec

Title: US-09-987-455-7

Perfect score: 1065

Sequence: 1 tctgagggaacagtgactg.....gtgacaacatcgaccctga 1065

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N\_Geneseq\_19Jun03.\*

1: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.\*  
2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.\*  
3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.\*  
4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.\*  
5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.\*  
6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.\*  
7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.\*  
8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.\*  
9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.\*  
10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.\*  
11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.\*  
12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.\*  
13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.\*  
14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.\*  
15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.\*  
16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.\*  
17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.\*  
18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.\*  
19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.\*  
20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.\*  
21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.\*  
22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.\*  
23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.\*  
24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.\*  
25: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1065	100.0	1065	AAD40616	Human kringle 2 se
2	1065	100.0	1068	AAN91120	Sequence of coding
3	1065	100.0	1170	AAN91135	Sequence of coding
4	1065	100.0	1314	AAN91123	Sequence of coding
5	1065	100.0	1419	AAN91133	Sequence of coding
6	1065	100.0	1680	AAT27588	Novel plasminogen
7	1065	100.0	1680	AAT27587	Novel plasminogen
8	1065	100.0	1689	AAN82178	Tissue plasminogen

9	1065	100.0	1780	11	AAQ01358	Sequence encoding
10	1065	100.0	2100	10	AAN91119	Sequence of native
11	1065	100.0	2100	11	AAQ05532	Plasmin pST112 enc
12	1065	100.0	2162	16	AAQ87370	Human tPA clone.
13	1065	100.0	2162	19	AAV37294	Human tissue plas
14	1065	100.0	2547	7	AAN60659	Sequence encoding
15	1065	100.0	2560	10	AAN90542	Plasmin PKG12 cont
16	1065	100.0	7533	8	AAN70491	Entire sequence of
17	1065	100.0	7533	9	AAN81970	Plasmin pEMpi-tPA.
18	1063.4	99.8	1614	11	AAQ03581	Glycosylated plas
19	1063.4	99.8	1614	12	AAQ11550	Sequence encoding
20	1063.4	99.8	1659	12	AAQ11551	Sequence encoding
21	1063.4	99.8	1689	9	AAN82179	Tissue plasminogen
22	1063.4	99.8	1689	11	AAN81090	Mutated t-PA analo
23	1063.4	99.8	1689	11	AAQ04903	Part of tPA024 gen
24	1063.4	99.8	1689	11	AAQ04904	Part of tPA023 gen
25	1063.4	99.8	1689	11	AAQ05177	Sequence encoding
26	1063.4	99.8	1689	12	AAQ13379	T-PA67+ mutant wit
27	1063.4	99.8	1689	24	ABK52546	Human tissue plas
28	1063.4	99.8	1740	17	AAT27585	Full-length tissue
29	1063.4	99.8	1955	21	AAZ31998	Human tissue plas
30	1063.4	99.8	1989	7	AAN60700	Sequence encoding
31	1063.4	99.8	2091	12	AAQ12073	T-PA variant contg
32	1063.4	99.8	2099	11	AAQ05534	Plasmin plgNTPA
33	1063.4	99.8	2100	12	AAQ12072	T-PA with -ve char
34	1063.4	99.8	2170	8	AAN70248	Sequence encoding
35	1063.4	99.8	2170	16	AAQ86576	Human tissue plas
36	1063.4	99.8	2249	11	AAQ05535	Plasmin plgN delta
37	1063.4	99.8	2457	7	AAN60146	cDNA sequence enco
38	1063.4	99.8	2509	10	AAN90508	cDNA of human tiss
39	1063.4	99.8	2519	24	ABK52545	Human tissue plas
40	1063.4	99.8	2519	24	ABL59544	Human tissue type
41	1063.4	99.8	2547	4	AAN30001	Sequence encoding
42	1063.4	99.8	7360	17	AAT15930	DHFR/intron (Wtras
43	1062	99.7	1128	24	AAD40614	OmpA-K2S fusion pr
44	1062	99.7	1341	11	AAQ06762	Sequence encoding
45	1061.8	99.7	1068	10	AAN91122	Sequence of coding

#### ALIGNMENTS

##### RESULT 1

AAD40616	ID	AAD40616	standard; DNA; 1065 BP.
XX	AC	AAD40616;	
XX	DT	30-OCT-2002	(first entry)
XX	DE	Human kringle 2	serine protease (K2S) DNA.
XX	KW	Tissue plasminogen	activator; tPA; kringle 2 serine protease; stroke;
XX	KW	K2S; cardiac	infarction; acute myocardial infarction; pulmonary embolism;
XX	KW	artery occlusion;	deep vein thrombosis; blood clotting; thrombolytic;
XX	OS	cerebroprotective;	cardiant; ompA; enzyme; human; gene; ds.
XX	FT	Homo sapiens.	
XX	Key.	Location/Qualifiers	
FT	CDS	1..1065	
FT	FT	/*tag= a	
FT	FT	/product= "Human K2S protein"	
XX	XX	WO200240650-A2.	
XX	PN		
XX	PD	23-MAY-2002.	
XX	PP	07-NOV-2001; 2001WO-EPI2857.	
XX	XX	14-NOV-2000; 2000GB-0027779.	
XX	PR		
XX	XX	(BOEH ) BOEHRINGER INGELHEIM INT GMBH.	

XX Werner R, Goetz F, Tayapiwatana C, Manosroi J, Manosroi A;  
 XX WPI; 2002-519376/55.  
 DR P-PSDB; AAE25036.  
 XX  
 PT Producing active, correctly folded recombinant tissue plasminogen  
 PT activator, Kingle 2 serine protease in prokaryotic cells by expressing  
 PT the protein-encoding DNA operably linked to DNA coding for signal  
 PT peptide OmpA -  
 XX  
 PS Claim 12; Page 31; 80pp; English.  
 XX  
 CC The present invention relates to a method of producing extracellularly  
 CC secreted, active, correctly folded, recombinant tissue plasminogen  
 CC activator (tPA), Kingle 2 serine protease molecule (K2S) or their  
 CC variants in prokaryotic cells by expressing the protein-encoding DNA  
 CC operably linked to DNA coding for signal peptide OmpA. The method is  
 CC useful for producing recombinant DNA-derived tPA, K2S or their variants.  
 CC Sequences of the invention are useful for manufacturing a medicament  
 CC for treating stroke, cardiac infarction, acute myocardial infarction,  
 CC pulmonary embolism, any artery occlusion such as intracranial artery  
 CC occlusion (e.g., arteries supplying the brain), peripherally occluded  
 CC arteries, coronary artery occlusion, deep vein thrombosis or related  
 CC diseases associated with unwanted blood clotting. The present sequence  
 CC is human K2S DNA.  
 XX  
 SQ Sequence 1065 BP; 225 A; 314 C; 312 G; 214 T; 0 other;

Query Match 100.0%; Score 1065; DB 24; Length 1065;  
 Best Local Similarity 100.0%; Pred. No. 9.7e-237;  
 Matches 1065; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCTGGGAAACAGTACTGTCTTGGGAATGGTCCAGCTACCGTGGCAGCAGC 60  
 DB 1 TCTGGGAAACAGTACTGTCTTGGGAATGGTCCAGCTACCGTGGCAGCAGC 60  
 QY 61 CTCACCGAGTCGGGTGCTCTGCTCCGCTGGAAATCCATGATCCTGATAGGCAAGTT 120  
 DB 61 CTCACCGAGTCGGGTGCTCTGCTCCGCTGGAAATCCATGATCCTGATAGGCAAGTT 120  
 QY 121 TACACAGACAGAAACCCAGTCCCGAGCACTGGGCTGGGCAACATAATTAAGTCGGG 180  
 DB 121 TACACAGACAGAAACCCAGTCCCGAGCACTGGGCTGGGCAACATAATTAAGTCGGG 180  
 QY 181 AATCTGATGGGATGCCAAGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240  
 DB 181 AATCTGATGGGATGCCAAGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240  
 QY 241 GAGTACTGTGATGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300  
 DB 241 GAGTACTGTGATGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300  
 QY 301 TTTGCAATCAAGAGGGCTCTTGGCGACATCGCTCCCAACCCCTGGCAGGCTGCCATC 360  
 DB 301 TTTGCAATCAAGAGGGCTCTTGGCGACATCGCTCCCAACCCCTGGCAGGCTGCCATC 360  
 QY 361 TTTGCAAGCAGCAGGAGTGGCCGGAGAGCGGTCTCTGTCGGGGGATACTCATCAGC 420  
 DB 361 TTTGCAAGCAGCAGGAGTGGCCGGAGAGCGGTCTCTGTCGGGGGATACTCATCAGC 420  
 QY 421 TCCTGCTGATTCCTCTGTCGGCCACCTGCTTCAGAGGAGGTTTCGCCCCACACCTG 480  
 DB 421 TCCTGCTGATTCCTCTGTCGGCCACCTGCTTCAGAGGAGGTTTCGCCCCACACCTG 480  
 QY 481 ACGGTGATCTTTGGCAGAACATACCGGTGCTCCCTGGCAGGAGGAGCAAAATTTGAA 540  
 DB 481 ACGGTGATCTTTGGCAGAACATACCGGTGCTCCCTGGCAGGAGGAGCAAAATTTGAA 540  
 QY 541 GTCGAAATACATTTGTCATAAGGAATTCATGATGACCTTACGCAATGACATTCG 600  
 DB 541 GTCGAAATACATTTGTCATAAGGAATTCATGATGACCTTACGCAATGACATTCG 600

QY 601 CTGTGCAGCTGAATCGGATTCGTCCCGTGTGCTCCAGAGAGCAGCGTGGTCCGCACT 660  
 DB 601 CTGTGCAGCTGAATCGGATTCGTCCCGTGTGCTCCAGAGAGCAGCGTGGTCCGCACT 660  
 QY 661 GTGTGCCCTTCCCGCGGAGACCTGTGAGCTCCGAGCTGGACGAGTGTGAGCTCTCCGGC 720  
 DB 661 GTGTGCCCTTCCCGCGGAGACCTGTGAGCTCCGAGCTGGACGAGTGTGAGCTCTCCGGC 720  
 QY 721 TAGCGCAAGCATGAGGCTTGTCTCTTCTTATTCGGAGCGGCTGAAGGAGGCTCATGTC 780  
 DB 721 TAGCGCAAGCATGAGGCTTGTCTCTTCTTATTCGGAGCGGCTGAAGGAGGCTCATGTC 780  
 QY 781 AGATGTACCATCCATCCAGCCGCTGCATCATCAACATTTACTTAACAGAACAGTCAACGAC 840  
 DB 781 AGATGTACCATCCAGCCGCTGCATCATCAACATTTACTTAACAGAACAGTCAACGAC 840  
 QY 841 AACATGCTGTGTGCTGGAGACACTCGAGCGGGGGCCCCAGGCAAACTTGCACGAGGCC 900  
 DB 841 AACATGCTGTGTGCTGGAGACACTCGAGCGGGGGCCCCAGGCAAACTTGCACGAGGCC 900  
 QY 901 TGGCAGGCGGATTCGGAGGCGCCCTGCTGCTGCTGAACGATGGCCGCTGCTTGGTG 960  
 DB 901 TGGCAGGCGGATTCGGAGGCGCCCTGCTGCTGCTGAACGATGGCCGCTGCTTGGTG 960  
 QY 961 GGCATCATCACTGCTGGGCTGGGCTGTGGACAGAGGATGTCCCGGTGTGTACACAAAG 1020  
 DB 961 GGCATCATCACTGCTGGGCTGGGCTGTGGACAGAGGATGTCCCGGTGTGTACACAAAG 1020  
 QY 1021 GTTACCAACTACCTAGACTGGATTCGTGACAAACATGCGACCGTGA 1065  
 DB 1021 GTTACCAACTACCTAGACTGGATTCGTGACAAACATGCGACCGTGA 1065

RESULT 2  
 AAN91120  
 ID AAN91120 standard; DNA; 1068 BP.  
 XX  
 AC AAN91120;  
 XX  
 AC AAN91120;  
 DT 25-MAR-2003 (updated)  
 DT 03-OCT-2002 (updated)  
 DT 18-JUN-1990 (first entry)  
 XX  
 DE Sequence of coding region in plasmid pStII.2.  
 XX  
 KW Tissue plasminogen activator; tPA; thrombolytic agent;  
 KW plasminogen; vascular diseases.  
 XX  
 OS Synthetic.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 1..1065  
 FT /\*tag= a  
 XX  
 EP302456-A.  
 XX  
 FN 08-FEB-1989.  
 PD  
 XX 02-AUG-1988; 88EP-0112569.  
 PF  
 XX 03-AUG-1987; 87GB-0018298.  
 PR 26-OCT-1987; 87GB-0025052.  
 PR 13-NOV-1987; 87GB-0026683.  
 XX  
 (FUJI ) FUJISAWA PHARM CO LTD.  
 FA  
 XX Niwa M, Saito Y, Sasaki H, Hayashi M, Notani J, Kobayashi M;  
 PI WPI; 1989-040625/06.  
 XX P-PSDB; AAP94407.  
 DR  
 XX New tissue plasminogen activator -  
 PT comprising finger and growth factor domains lacking tPA for



```
QY 121 TACACAGCAGAACCCAGTCCAGGCACTGGCCCTGGGCAACATAATTTACTGCGG 180
DB |||||
QY 226 TACACAGCAGAACCCAGTCCAGGCACTGGCCCTGGGCAACATAATTTACTGCGG 285
DB |||||
QY 181 AATCCTGATGGGATGCCAAGCCCTGTGTCACATGCTGTAAGAACCCGAGCTGACGTGG 240
DB |||||
QY 286 AATCCTGATGGGATGCCAAGCCCTGTGTCACATGCTGTAAGAACCCGAGCTGACGTGG 345
DB |||||
QY 241 GAGTACTGTGATGTGCGCTCTCTGCTCCACCTGCGGCTGAGACAGTACAGCCGCTCAG 300
DB |||||
QY 346 GAGTACTGTGATGTGCGCTCTCTGCTCCACCTGCGGCTGAGACAGTACAGCCGCTCAG 405
DB |||||
QY 301 TTTGCGCATCAAGAGGAGGCTCTTTCGCGACATCGCTCCACCCCTGGCAGGCTGCCATC 360
DB |||||
QY 406 TTTGCGCATCAAGAGGAGGCTCTTTCGCGACATCGCTCCACCCCTGGCAGGCTGCCATC 465
DB |||||
QY 361 TTTGCGCAAGCAGAGGAGTCCGCGGAGAGCGGTTCTGTCGGGGGATACTCATCAGC 420
DB |||||
QY 466 TTTGCGCAAGCAGAGGAGTCCGCGGAGAGCGGTTCTGTCGGGGGATACTCATCAGC 525
DB |||||
QY 421 TCCTGCTGGATTCCTCTGCGGCGGCACTGCTTCAGGAGAGGTTTCGCGCGCAGCAGCTG 480
DB |||||
QY 526 TCCTGCTGGATTCCTCTGCGGCGGCACTGCTTCAGGAGAGGTTTCGCGCGCAGCAGCTG 585
DB |||||
QY 481 ACGGTGATCTTGGGCGAGAACATACCGGCTGTGCTCCCTGGCGAGGAGGACAGAAATTTGAA 540
DB |||||
QY 586 ACGGTGATCTTGGGCGAGAACATACCGGCTGTGCTCCCTGGCGAGGAGGACAGAAATTTGAA 645
DB |||||
QY 541 GTCGAAAATACATTTGTCATAAGGAATTCATGATGACACTTACGACAAATGACATTCGC 600
DB |||||
QY 646 GTCGAAAATACATTTGTCATAAGGAATTCATGATGACACTTACGACAAATGACATTCGC 705
DB |||||
QY 601 CTGCTGACAGTGAATTCGGATTCGTCGCTGTGTCGCGGAGGAGGAGGCTGTCGCACT 660
DB |||||
QY 706 CTGCTGACAGTGAATTCGGATTCGTCGCTGTGTCGCGGAGGAGGAGGCTGTCGCACT 765
DB |||||
QY 661 GTGTGCTCTTCCCGCGGAGCTGCGAGCTGCGGAGTGGAGTGGAGTCTCTCGGC 720
DB |||||
QY 766 GTGTGCTCTTCCCGCGGAGCTGCGAGCTGCGGAGTGGAGTGGAGTCTCTCGGC 825
DB |||||
QY 721 TAGCGCAAGCATAGGCGCTTGTCTCTTTCTATTGCGAGCGGCTGAAGAGGCTCATGTC 780
DB |||||
QY 826 TAGCGCAAGCATAGGCGCTTGTCTCTTTCTATTGCGAGCGGCTGAAGAGGCTCATGTC 885
DB |||||
QY 781 AGACTGTACCATCCAGCGCTGACATCAACATTTACTTAAACAGAACAGTCAACGAC 840
DB |||||
QY 886 AGACTGTACCATCCAGCGCTGACATCAACATTTACTTAAACAGAACAGTCAACGAC 945
DB |||||
QY 841 AACATGTGTGTGCTGGAGACACTCGGAGCGGCGGCGGCGGCAAACTTGCACGCGCC 900
DB |||||
QY 946 AACATGTGTGTGCTGGAGACACTCGGAGCGGCGGCGGCGGCAAACTTGCACGCGCC 1005
DB |||||
QY 901 TGCCAGGCGGATTCGGAGGCGGCTGCTGTGTGTGTAACGATGCGCGCATGCTTTGGTG 960
DB |||||
QY 1006 TGCCAGGCGGATTCGGAGGCGGCTGCTGTGTGTGTAACGATGCGCGCATGCTTTGGTG 1065
DB |||||
QY 961 GGCATCATAGCTGGGCGCTGGGCTGTGACAGAGGATGTCCGGGTGTGTACAAAG 1020
DB |||||
QY 1066 GGCATCATAGCTGGGCGCTGGGCTGTGACAGAGGATGTCCGGGTGTGTACAAAG 1125
DB |||||
QY 1021 GTTACCAACTACCTAGATCGGATTCGTGACAAACATGCGACCGTGA 1065
DB |||||
QY 1126 GTTACCAACTACCTAGATCGGATTCGTGACAAACATGCGACCGTGA 1170
DB |||||
```

## RESULT 4

AA91123

ID AA91123 standard; DNA; 1314 BP.

XX

AC AA91123;

XX

DT 25-MAR-2003 (updated)

```
DT 03-OCT-2002 (updated)
DT 18-JUN-1990 (first entry)
XX
XX Sequence of coding region in plasmid pTQkPA delta trp.
XX
XX Tissue plasminogen activator; tPA; thrombolytic agent;
XX plasminogen; vascular diseases.
XX
XX Synthetic.
XX
XX Key Location/Qualifiers
XX CDS 1..1065
XX /*tag= a
XX
XX EP302456-A.
XX
XX 08-FEB-1989.
XX
XX 02-AUG-1988; 88EP-0112569.
XX
XX 03-AUG-1987; 87GB-0018298.
XX 26-OCT-1987; 87GB-0025052.
XX 13-NOV-1987; 87GB-0026683.
XX
XX (FUJI ) FUJISAWA PHARM CO LTD.
XX
XX Niwa M, Saito Y, Sasaki H, Hayashi M, Notani J, Kobayashi M;
XX
XX WPI; 1989-040625/06.
XX P-PSDB; AAP94409.
XX
XX New tissue plasminogen activator -
XX comprising finger and growth factor domains lacking tPA for
XX longer half-life and stronger thrombolytic activity.
XX
XX Disclosure; Page ?; 69pp; English.
XX
XX
XX (Updated on 03-OCT-2002 to add missing OS field.)
XX (Updated on 25-MAR-2003 to correct PA field.)
XX
XX Sequence 1314 BP; 286 A; 386 C; 393 G; 249 T; 0 other;
```

Query Match 100.0%; Score 1065; DB 10; Length 1314;  
Best Local Similarity 100.0%; Pred. No. 1e-236;  
Matches 1065; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 TCTGAGGAAACAGTGACTGCTACTTTGGGAATGGGTACGCTTACCGTGGCACGACAGC 60
DB |||||
QY 250 TCTGAGGAAACAGTGACTGCTACTTTGGGAATGGGTACGCTTACCGTGGCACGACAGC 309
DB |||||
QY 61 CTACACGAGTGGGTGCTCTGCTCCCTCCGCTGGGAATTCATGATCCTGTAGGCAAGGTT 120
DB |||||
QY 310 CTACACGAGTGGGTGCTCTGCTCCCTCCGCTGGGAATTCATGATCCTGTAGGCAAGGTT 369
DB |||||
QY 121 TACACAGCAGAACCCAGTCCAGGCACTGGGCGCTGGGCAACATAATTTACTGCGG 180
DB |||||
QY 370 TACACAGCAGAACCCAGTCCAGGCACTGGGCGCTGGGCAACATAATTTACTGCGG 429
DB |||||
QY 181 AATCCTGATGGGATGCCAAGCCCTGTGTCACATGCTGTAAGAACCCGAGCTGACGTGG 240
DB |||||
QY 430 AATCCTGATGGGATGCCAAGCCCTGTGTCACATGCTGTAAGAACCCGAGCTGACGTGG 489
DB |||||
QY 241 GAGTACTGTGATGTGCGCTCTCTGCTCCACCTGCGGCTGAGACAGTACAGCCGCTCAG 300
DB |||||
QY 490 GAGTACTGTGATGTGCGCTCTCTGCTCCACCTGCGGCTGAGACAGTACAGCCGCTCAG 549
DB |||||
QY 301 TTTGCGCATCAAGAGGAGGCTCTTTCGCGACATCGCTCCACCCCTGGCAGGCTGCCATC 360
DB |||||
QY 550 TTTGCGCATCAAGAGGAGGCTCTTTCGCGACATCGCTCCACCCCTGGCAGGCTGCCATC 609
DB |||||
QY 361 TTTGCGCAAGCAGAGGAGTCCGCGGAGAGCGGTTCTGTCGGGGGATACTCATCAGC 420
DB |||||
QY 610 TTTGCGCAAGCAGAGGAGTCCGCGGAGAGCGGTTCTGTCGGGGGATACTCATCAGC 669
DB |||||
```

QY 421 TCCTGCTGGATTCTCTCTGCGGCCCACTGCTTTCAGGAGAGGTTTCCGCCCCACACCTG 480  
 Db 670 TCCTGCTGGATTCTCTCTGCGGCCCACTGCTTTCAGGAGAGGTTTCCGCCCCACACCTG 729  
 QY 481 ACGGTGATCTTGGGAGAGACATACCGGTGCTCCCTGCGGAGGAGGACAGAAATTTGAA 540  
 Db 730 ACGGTGATCTTGGGAGAGACATACCGGTGCTCCCTGCGGAGGAGGACAGAAATTTGAA 789  
 QY 541 GTCGAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAATGACATTCGCG 600  
 Db 790 GTCGAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAATGACATTCGCG 849  
 QY 601 CTGCTGAGCTGAAATCGAATTCGTCGCTGTGTCGCGAGAGAGCGTGGTCGCACT 660  
 Db 850 CTGCTGAGCTGAAATCGAATTCGTCGCTGTGTCGCGAGAGAGCGTGGTCGCACT 909  
 QY 661 GTGTGCTTCCCGCGGAGCTGAGCTGCGGAGCTGCGGAGCTGCGGAGCTGCGGAGCTGCGG 720  
 Db 910 GTGTGCTTCCCGCGGAGCTGAGCTGCGGAGCTGCGGAGCTGCGGAGCTGCGGAGCTGCGG 969  
 QY 721 TACGCAAGCATGAGGCTTGTCTCTTCTATTTCGAGCGGCTGAAGAGGCTCATGTC 780  
 Db 970 TACGCAAGCATGAGGCTTGTCTCTTCTATTTCGAGCGGCTGAAGAGGCTCATGTC 1029  
 QY 781 AGACTGTACCATTCAGCGCTGCAATCAACATTTACTTAAACAGAACAGTCAACGAC 840  
 Db 1030 AGACTGTACCATTCAGCGCTGCAATCAACATTTACTTAAACAGAACAGTCAACGAC 1089  
 QY 841 AACATGCTGTGCTGGAGACTCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 900  
 Db 1090 AACATGCTGTGCTGGAGACTCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1149  
 QY 901 TGCCAGGCGGATTCGCGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 960  
 Db 1150 TGCCAGGCGGATTCGCGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1209  
 QY 961 GGCATCATAGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCT 1020  
 Db 1210 GGCATCATAGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCT 1269  
 QY 1021 GTTACCACTACCTAGACTGGATTCGTCACACATGCGACCGTGA 1065  
 Db 1270 GTTACCACTACCTAGACTGGATTCGTCACACATGCGACCGTGA 1314

## RESULT 5

AAN91133

ID AAN91133 standard; DNA; 1419 BP.

XX AC AAN91133;

XX DT 25-MAR-2003 (updated)

XX DT 03-OCT-2002 (updated)

XX DT 18-JUN-1990 (first entry)

XX DE Sequence of coding region in plasmid pmQk112.

XX KW Tissue plasminogen activator; tPA; thrombolytic agent;

XX KW plasminogen; vascular diseases.

XX OS Synthetic.

XX XX Key

XX FT Location/Qualifiers

XX FT 1..1065

XX FT /\*tag= a

XX XX EP302456-A.

XX XX 08-FEB-1989.

XX XX 02-AUG-1988; 88EP-0112569.

XX XX

PR 03-AUG-1987; 87GB-0018298.  
 PR 26-OCT-1987; 87GB-0025052.  
 PR 13-NOV-1987; 87GB-0026683.  
 XX  
 PA (FUJI ) FUJISAWA PHARM CO LTD.  
 XX  
 XX Niwa M, Saito Y, Sasaki H, Hayashi M, Notani J, Kobayashi M;  
 XX WPI; 1989-040625/06.  
 DR P-PSDB; AAP94416.  
 XX  
 XX New tissue plasminogen activator -  
 PT comprising finger and growth factor domains lacking tPA for  
 PT longer half-life and stronger thrombolytic activity.  
 XX  
 XX Disclosure; Page ?; 68pp; English.  
 XX  
 CC (Updated on 03-OCT-2002 to add missing OS field.)  
 CC (Updated on 25-MAR-2003 to correct PA field.)  
 XX  
 SQ Sequence 1419 BP; 309 A; 413 C; 426 G; 271 T; 0 other;

Query Match 100.0%; Score 1065; DB 10; Length 1419;  
 Best Local Similarity 100.0%; Pred. No. 1e-236;  
 Matches 1065; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 TCTGAGGAAACAGTGAAGTCTTCTGCTTGGGAATGGGTGAGCTTACCGTGGCAGCAGCAGC 60  
 Db 355 TCTGAGGAAACAGTGAAGTCTTCTGCTTGGGAATGGGTGAGCTTACCGTGGCAGCAGCAGC 414  
 QY 61 CTACACGAGTGGGTGCTTCTGCTTGGGAATGGGTGAGCTTACCGTGGCAGCAGCAGCAGC 120  
 Db 415 CTACACGAGTGGGTGCTTCTGCTTGGGAATGGGTGAGCTTACCGTGGCAGCAGCAGCAGC 474  
 QY 121 TACACACGAGTGGGTGCTTCTGCTTGGGAATGGGTGAGCTTACCGTGGCAGCAGCAGCAGC 180  
 Db 475 TACACACGAGTGGGTGCTTCTGCTTGGGAATGGGTGAGCTTACCGTGGCAGCAGCAGCAGC 534  
 QY 181 AATCCTGATGGGATGCAAGCCCTGCTGCTGCAACCTGCTGCAACCTGCTGCAACCTGCTGCA 240  
 Db 535 AATCCTGATGGGATGCAAGCCCTGCTGCTGCAACCTGCTGCAACCTGCTGCAACCTGCTGCA 594  
 QY 241 GAGTACTGATGTCCTTCTGCTTGGGAATGGGTGAGCTTACCGTGGCAGCAGCAGCAGCAGC 300  
 Db 595 GAGTACTGATGTCCTTCTGCTTGGGAATGGGTGAGCTTACCGTGGCAGCAGCAGCAGCAGC 654  
 QY 301 TTTGCGCATCAAGAGGAGGCTTCTGCTGCGGAGATGCTGCTGCGGAGATGCTGCTGCGGAG 360  
 Db 655 TTTGCGCATCAAGAGGAGGCTTCTGCTGCGGAGATGCTGCTGCGGAGATGCTGCTGCGGAG 714  
 QY 361 TTTGCGCATCAAGAGGAGGCTTCTGCTGCGGAGATGCTGCTGCGGAGATGCTGCTGCGGAG 420  
 Db 715 TTTGCGCATCAAGAGGAGGCTTCTGCTGCGGAGATGCTGCTGCGGAGATGCTGCTGCGGAG 774  
 QY 421 TCCTGCTGATTCCTTCTGCTGCGGAGATGCTGCTGCGGAGATGCTGCTGCGGAGATGCTGCT 480  
 Db 775 TCCTGCTGATTCCTTCTGCTGCGGAGATGCTGCTGCGGAGATGCTGCTGCGGAGATGCTGCT 834  
 QY 481 ACGGTGATCTTGGGAGAGACATACCGGTGCTCCCTGCGGAGGAGGAGGAGGAGGAGGAGGAG 540  
 Db 835 ACGGTGATCTTGGGAGAGACATACCGGTGCTCCCTGCGGAGGAGGAGGAGGAGGAGGAGGAG 894  
 QY 541 GTCGAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAATGACATTCGCG 600  
 Db 895 GTCGAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAATGACATTCGCG 954  
 QY 601 CTGCTGAGCTGAAATCGAATTCGTCGCTGTGTCGCGAGAGAGCGTGGTCGCACT 660  
 Db 955 CTGCTGAGCTGAAATCGAATTCGTCGCTGTGTCGCGAGAGAGCGTGGTCGCACT 1014  
 QY 661 GTGTGCTTCCCGCGGAGCTGAGCTGCGGAGCTGCGGAGCTGCGGAGCTGCGGAGCTGCGGAG 720  
 Db 1015 GTGTGCTTCCCGCGGAGCTGAGCTGCGGAGCTGCGGAGCTGCGGAGCTGCGGAGCTTCTCGGC 1074





Db 1456 AACATGCTGTGTGGAGACACTCGAGCGCGCGGCCCGCCAGGCAAACTTGCACGACGCC 1515  
Qy 901 TGCCAGGCGGATTTCCGGAGGCCCCCTGGTGTCTGAAACGATGGCCGATGACTTTTGGTG 960  
Db 1516 TGCCAGGCGGATTTCCGGAGGCCCCCTGGTGTCTGAAACGATGGCCGATGACTTTGGTG 1575  
Qy 961 GGCATCATCAGCTGGGGCTGGGCTGTGGACAGAGATGTCCCGGGTGTGTACACAAAG 1020  
Db 1576 GGCATCATCAGCTGGGGCTGGGCTGTGGACAGAGATGTCCCGGGTGTGTACACAAAG 1635  
Qy 1021 GTTACCACTACCTAGACTGGATTGGTGCACACATGCGACCGTGA 1065  
Db 1636 GTTACCACTACCTAGACTGGATTGGTGCACACATGCGACCGTGA 1680

RESULT 7

AAT27587

ID AAT27587 standard; DNA; 1680 BP.

XX AC AAT27587;

DT 25-MAR-2003 (updated)

XX DT 06-AUG-1996 (first entry)

XX DE Novel plasminogen activator DNA.

XX Tissue plasminogen activator; tPA; alpha2-plasmin inhibitor;  
KW fibrinolytic; thrombolytic; fibrin; thrombosis; blood clotting;  
KW protein engineering; kringle; ss.  
XX Synthetic.

OS Key

XX Location/Qualifiers

FT sig\_peptide 1..105

FT mat\_peptide 106..1677

FT /tag= a

FT /tag= b

XX US5504001-A.

XX 02-APR-1996.

XX 06-JUN-1994; 94US-0254485.

XX 25-NOV-1987; 87US-0125629.

XX 28-JAN-1992; 92US-0827587.

XX 06-JUN-1994; 94US-0254485.

XX (ZYMO ) ZYMOGENETICS INC.

XX Foster DC;

XX WPI; 1996-187699/19.

XX P-PSDB; AAR96222.

XX Hybrid plasminogen activator comprises human tPA activator and

XX N-terminal crosslinking domain from alpha2-plasmin inhibitor

XX useful to treat thrombosis and image blood clots

XX Example 3; Fig 10; 35pp; English.

XX A DNA construct (AAT27587) codes for a novel plasminogen  
XX activator (AAR96223) in which the kringle K1 domain of plasminogen  
XX (see also AAR96221) is mutated to substitute asparagine for aspartic  
XX acid at position 5, and replaces the native K1 domain of tissue  
XX plasminogen activator (tPA) (AAR96220). It was obtd. by mutagenesis  
XX of the sequence in vector ZEM99-8000 (see also AAT27587). The novel  
XX protein can be expressed in Escherichia coli RRI/Zem99-8010 cells  
XX (FERM P-9315). Novel plasminogen activators are produced that show  
XX increased clot lysing specificity or plasma half-life.  
XX (Updated on 25-MAR-2003 to correct PF field.)

XX

SQ Sequence 1680 BP; 386 A; 471 C; 480 G; 343 T; 0 other;  
Query Match 100.0%; Score 1065; DB 17; Length 1680;  
Best Local Similarity 100.0%; Pred. No. 1.1e-236;  
Matches 1065; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 TCTGAGGAAACAGTGAAGTCTCTCTGCTCCGCTGGAAATTCATGATCCCTGATAGGCAAGTT 120  
Db 616 TCTGAGGAAACAGTGAAGTCTCTCTGCTCCGCTGGAAATTCATGATCCCTGATAGGCAAGTT 735  
Qy 61 CTACCCAGTGGGTGCTCTCTGCTCCGCTGGAAATTCATGATCCCTGATAGGCAAGTT 120  
Db 676 CTCACCCAGTGGGTGCTCTCTGCTCCGCTGGAAATTCATGATCCCTGATAGGCAAGTT 735  
Qy 121 TACACAGCACAGAACCCAGTGGCCAGGCACTGGGCTGGGCAACATATTAATCTGCGG 180  
Db 736 TACACAGCACAGAACCCAGTGGCCAGGCACTGGGCTGGGCAACATATTAATCTGCGG 795  
Qy 181 AATCCTGATGGGATGCCAAGCCCTGTGTGCCACCTGTCTGAAGAACCCGAGGCTGACGTGG 240  
Db 796 AATCCTGATGGGATGCCAAGCCCTGTGTGCCACCTGTCTGAAGAACCCGAGGCTGACGTGG 855  
Qy 241 GAGTACTGTGATGTCCTCTCTGCTCCACCTGCGGCTTGAGACAGTACAGCCGCTCAG 300  
Db 856 GAGTACTGTGATGTCCTCTCTGCTCCACCTGCGGCTTGAGACAGTACAGCCGCTCAG 915  
Qy 301 TTTGCGATCAAGAGAGGGCTCTTCCGCGACATCCCTCCACCCCTGGCAGGCTGCCATC 360  
Db 916 TTTGCGATCAAGAGAGGGCTCTTCCGCGACATCCCTCCACCCCTGGCAGGCTGCCATC 975  
Qy 361 TTTGCCAAGCACAGGAGGTGCGCCGAGAGCGGTTCCTGTGCGGGGGCATACTCATCAGC 420  
Db 976 TTTGCCAAGCACAGGAGGTGCGCCGAGAGCGGTTCCTGTGCGGGGGCATACTCATCAGC 1035  
Qy 421 TCCTGTGGAATTCCTCTGCGGCGCCACTGCTTCAGGAGAGGTTTCGCCGCCACCACTG 480  
Db 1036 TCCTGTGGAATTCCTCTGCGGCGCCACTGCTTCAGGAGAGGTTTCGCCGCCACCACTG 1095  
Qy 481 ACGGTGATCTTGGGCAAGACATACCGGCTGGTCCCTGCGAGAGGAGGAGCAAAATTTGAA 540  
Db 1096 ACGGTGATCTTGGGCAAGACATACCGGCTGGTCCCTGCGAGGAGGAGGAGCAAAATTTGAA 1155  
Qy 541 GTCGAAAAATACATTTGTCCATAAGGAATTCGATGACACTTACGACAAATGACATTCGCG 600  
Db 1156 GTCGAAAAATACATTTGTCCATAAGGAATTCGATGACACTTACGACAAATGACATTCGCG 1215  
Qy 601 CTGCTGACAGTGAATTCGGAATTCGCTCCGCTGTCGCCAGGAGAGCAGGTGTCGCACT 660  
Db 1216 CTGCTGACAGTGAATTCGGAATTCGCTCCGCTGTCGCCAGGAGAGCAGGTGTCGCACT 1275  
Qy 661 GTGTGCTTCCCGCGGAGACCTGACAGTCCGAGCTGGAGCGAGTGTGAGCTCTCCGCGC 720  
Db 1276 GTGTGCTTCCCGCGGAGACCTGACAGTCCGAGCTGGAGCGAGTGTGAGCTCTCCGCGC 1335  
Qy 721 TACGGCAAGCATGAGGCTTTGTCTCTTCTTATTCGGAGCGGCTGAAGAGGCTCATGTC 780  
Db 1336 TACGGCAAGCATGAGGCTTTGTCTCTTCTTCTTATTCGGAGCGGCTGAAGAGGCTCATGTC 1395  
Qy 781 AGACTGTACCATCCAGCCGCTGACATCAACATTTACTTAAACAGACAGTCAACGAC 840  
Db 1396 AGACTGTACCATCCAGCCGCTGACATCAACATTTACTTAAACAGACAGTCAACGAC 1455  
Qy 841 AACATGCTGTGTGCTGGAGACACTCGGAGCGGGGGGCCCGAGGCAAACTTGCACGACGCC 900  
Db 1456 AACATGCTGTGTGCTGGAGACACTCGGAGCGGGGGGCCCGAGGCAAACTTGCACGACGCC 1515  
Qy 901 TGCCAGGCGGATTTCCGGAGGCCCCCTGGTGTCTGAAACGATGGCCGATGACTTTTGGTG 960  
Db 1516 TGCCAGGCGGATTTCCGGAGGCCCCCTGGTGTCTGAAACGATGGCCGATGACTTTGGTG 1575  
Qy 961 GGCATCATCAGCTGGGGCTGGGCTGTGGACAGAGATGTCCCGGGTGTGTACACAAAG 1020  
Db 1576 GGCATCATCAGCTGGGGCTGGGCTGTGGACAGAGATGTCCCGGGTGTGTACACAAAG 1635

QY 1021 GTTACCAACTACCTAGACTGGATTGCTGACAACTAGCGCGTGA 1065  
 |||||  
 DB 1636 GTTACCAACTACCTAGACTGGATTGCTGACAACTAGCGCGTGA 1680  
 |||||

## RESULT 8

AA001358  
 ID AA001358 standard; DNA; 1689 BP.

XX  
 AC AA001358;

XX 25-MAR-2003 (updated)

DT 10-MAR-2003 (updated)

DT 03-NOV-1990 (first entry)

DE Tissue plasminogen activator with S-119 substd for M.

XX Tissue plasminogen activator; tPA; thrombosis; N-glycosylation; ss.

XX Unidentified.

XX Key Location/Qualifiers  
 FT CDS 1..1689  
 FT /\*tag= a  
 FT /product=modified tPA

XX JP63230083-A.

XX 26-SEP-1988.

XX 20-MAR-1987; 87JP-0064339.

XX 20-MAR-1987; 87JP-0064339.

XX (EISA) EISAI CO LTD.

XX WPI; 1988-311961/44.

XX P-PSDB; AAP82581.

XX Modified tissue plasminogen activator - having glycine-183 and serine-186 residues substituted with serine and threonine.

XX Disclosure; Page 7; 16pp; Japanese.

XX One N-glycosylation site, i.e. NSS (117-119) is substituted with

CC NSM and the N-glycosylation is removed.

CC Plasmid encoding the modified tPA is 99-6300 and its transformant is

CC E.coli RRI-Zem 99-6300 (FERM P-9127).

CC This modified tPA, used to treat thrombosis, is of high quality and

CC has a longer half life period in blood.

CC See also AA001358-1987.

CC (Updated on 10-MAR-2003 to add missing OS field.)

CC (Updated on 25-MAR-2003 to correct PF field.)

CC (Updated on 25-MAR-2003 to correct PR field.)

CC (Updated on 25-MAR-2003 to correct PA field.)

XX Sequence 1689 BP; 376 A; 482 C; 504 G; 327 T; 0 other;

Query Match 100.0%; Score 1065; DB 9; Length 1689;

Best Local Similarity 100.0%; Pred. No. 1.1e-236;

Matches 1065; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCTGAGGAAACAGTACTGCTACTTGGGAATGGGTGAGCTACCGTGGGACGACAGC 60

DB 625 TCTGAGGAAACAGTACTGCTACTTGGGAATGGGTGAGCTACCGTGGGACGACAGC 684

QY 61 CTCACCGAGTCGGGTGCTCTCTGCTCCGCTGGGAATTCATGATCCTGATAGGCAAGTT 120

DB 685 CTCACCGAGTCGGGTGCTCTCTGCTCCGCTGGGAATTCATGATCCTGATAGGCAAGTT 744

QY 121 TACACAGCACAGAACCCAGTCGCCAGCACTGGGCTGGGCAACATAATTAATCTGCGG 180

|||||

DB 745 TACACAGCACAGAACCCAGTCGCCAGCACTGGGCTGGGCAACATAATTAATCTGCGG 804

QY 181 AATCTCATGGGATGCCAAGCCCTGCTGTCACCTGCTGAGAACCCGAGGCTGACGTTG 240

DB 805 AATCTCATGGGATGCCAAGCCCTGCTGTCACCTGCTGAGAACCCGAGGCTGACGTTG 864

QY 241 GAGTACTGTGATGTCCTCTCTGCTCCACCTGCGGCTGAGACAGTACAGCAGCCTCAG 300

DB 865 GAGTACTGTGATGTCCTCTCTGCTCCACCTGCGGCTGAGACAGTACAGCAGCCTCAG 924

QY 301 TTTCCGATCAAGAGAGGCTCTTCCGCGAGCATGCTCCCTCCACCTGCGAGGCTGCCATC 360

DB 925 TTTCCGATCAAGAGAGGCTCTTCCGCGAGCATGCTCCCTCCACCTGCGAGGCTGCCATC 984

QY 361 TTTGCCAAGCACAGGAGTCCCGGAGAGGGTTCTGTCGGGGGCGATCTCATCAGC 420

DB 985 TTTGCCAAGCACAGGAGTCCCGGAGAGGGTTCTGTCGGGGGCGATCTCATCAGC 1044

QY 421 TCCTGCTGGATTCTCTCTGCGCGCCACTGCTTCCAGGAGAGGTTTCCGCCCCCACCCTG 480

DB 1045 TCCTGCTGGATTCTCTCTGCGCGCCACTGCTTCCAGGAGAGGTTTCCGCCCCCACCCTG 1104

QY 481 ACGTGTATCTTGGGACAGAAATACCGGGTGGTCCCTGCGGAGGAGGAGCAAAATTTGAA 540

DB 1105 ACGTGTATCTTGGGACAGAAATACCGGGTGGTCCCTGCGGAGGAGGAGCAAAATTTGAA 1164

QY 541 GTCGAAAATACATTTGCTCCATAGGAATTCGATGATGACACTTACGACAAATGACATGCG 600

DB 1165 GTCGAAAATACATTTGCTCCATAGGAATTCGATGATGACACTTACGACAAATGACATGCG 1224

QY 601 CTGCTGACGTGAAATCGGATTTCCTCCGCTGTGCCAGGAGAGCAGCGTGGTCCGCACT 660

DB 1225 CTGCTGACGTGAAATCGGATTTCCTCCGCTGTGCCAGGAGAGCAGCGTGGTCCGCACT 1284

QY 661 GTGTGCTTCCCGCGGACCTGACGTGCGGACCTGCGAGAGTGTGAGCTCTCCGCGC 720

DB 1285 GTGTGCTTCCCGCGGACCTGACGTGCGGACCTGCGAGAGTGTGAGCTCTCCGCGC 1344

QY 721 TAGCGCAAGCATGAGGCTTGTCTCTTCTTATTCGAGCGGCTGAGGAGGCTCATGTC 780

DB 1345 TAGCGCAAGCATGAGGCTTGTCTCTTCTTATTCGAGCGGCTGAGGAGGCTCATGTC 1404

QY 781 AGCTGTACCCATCCAGCCGCTGCACATCAACATTTACTTTACAGAACAGTCACCGAC 840

DB 1405 AGCTGTACCCATCCAGCCGCTGCACATCAACATTTACTTTACAGAACAGTCACCGAC 1464

QY 841 AACATGCTGTGCTGAGACACTCGAGCGGCGGCCCCCAGGCAAACTTGTGACAGAGCC 900

DB 1465 AACATGCTGTGCTGAGACACTCGAGCGGCGGCCCCCAGGCAAACTTGTGACAGAGCC 1524

QY 901 TGCAGGCGGATTTCGGAGGCCCCCTGCTGTGCTGAACGATGCGCGCATGCTTTGTG 960

DB 1525 TGCAGGCGGATTTCGGAGGCCCCCTGCTGTGCTGAACGATGCGCGCATGCTTTGTG 1584

QY 961 GGATCATACAGTGGGCTGGGCTGTGGACAGAGGATGTCCGGGTGTGTACACAAAG 1020

DB 1585 GGATCATACAGTGGGCTGGGCTGTGGACAGAGGATGTCCGGGTGTGTACACAAAG 1644

QY 1021 GTTACCAACTACCTAGACTGGATTGCTGACAACTAGCGCGTGA 1065

DB 1645 GTTACCAACTACCTAGACTGGATTGCTGACAACTAGCGCGTGA 1689

## RESULT 9

AA001358

ID AA001358 standard; DNA; 1780 BP.

XX AA001358;

XX 25-MAR-2003 (updated)

DT 08-SEP-1990 (first entry)

XX Sequence encoding wild type tissue plasminogen activator (t-PA).

XX Wild type tissue plasminogen activator (t-PA); infarction treatment;  
 KW thrombosis treatment; embolism treatment.  
 XX Homo sapiens.  
 FH Key Location/Qualifiers  
 FT misc\_feature 1..123  
 FT /\*tag= a  
 FT /note="synthetic 5' adaptor"  
 FT 1738..1780  
 FT /\*tag= b  
 FT /note="synthetic 3' adaptor"  
 XX EP351246-A.  
 XX 17-JAN-1990.  
 XX 14-JUL-1989; 89EP-0307194.  
 XX 15-JUL-1988; 88DK-0003952.  
 XX (NOVO ) NOVO-NORDISK AS.  
 XX Petersen LC, Boel B;  
 XX WPI; 1990-016567/03.  
 XX P-PSDB; AAR04699.  
 XX New tissue plasminogen activator (t-PA) analogue -  
 PT with higher fibrin selectivity than native t-PA, useful for  
 PT treating infarction, thrombosis and embolism  
 XX Disclosure; Fig 5A-D; 24pp; English.  
 XX It is modified in the patent by replacing one or more codons specifying  
 CC a positively charged amino acid. The resulting analogues are inserted  
 CC into a replicable expression vector which is used to transform or  
 CC transfect a host cell which is grown to express a t-PA analogue. The  
 CC t-PA analogue is useful for the treatment of diseases or disorders  
 CC associated with the formation of thrombi in blood vessels e.g.  
 CC infarctions, thrombosis and embolism. The analogue, in the 1-chain form,  
 CC exhibits the properties of a proenzyme; however on plasmin-catalysed  
 CC cleavage of the one-chain form, the activity of the 2-chain form is fully  
 CC retained. Compared to native t-PA, it has a higher fibrin selectivity,  
 CC this results in a fibrinolytic agent with a higher fibrin selectivity,  
 CC as the fibrinogenolytic activity induced by the analogue is reduced  
 CC relative to that induced by native 1-chain t-PA.  
 CC (Updated on 25-MAR-2003 to correct PA field.)  
 XX SQ Sequence 1780 BP; 405 A; 513 C; 523 G; 339 T; 0 other;  
 Query Match 100.0%; Score 1065; DB 11; Length 1780;  
 Best Local Similarity 100.0%; Pred. No. 1.1e-236;  
 Matches 1065; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 TCTGAGGAACAGTACTGCTACTTTGGGAATGGGTGAGCTACCGTGGCAGCACGACAGC 60  
 DB 641 TCTGAGGAACAGTACTGCTACTTTGGGAATGGGTGAGCTACCGTGGCAGCACGACAGC 700  
 QY 61 CTCACCGAGTGGGTGCTCTCTGCTCCCGTGGGAATCCATGATCTCTGATAGGCAAGGTT 120  
 DB 701 CTCACCGAGTGGGTGCTCTCTGCTCCCGTGGGAATCCATGATCTCTGATAGGCAAGGTT 760  
 QY 121 TACACGACAGAACCCGAGCGCCAGCAGCTGGCTGGGCTGGGCAACATAATTAATCTGCGG 180  
 DB 761 TACACGACAGAACCCGAGCGCCAGCAGCTGGCTGGGCTGGGCAACATAATTAATCTGCGG 820  
 QY 181 AATCCTGATGGGATGCCAAGCCCTGTGTGCGACAGCTGTGGAAGAACCGAGCTGACCTGG 240  
 DB 821 AATCCTGATGGGATGCCAAGCCCTGTGTGCGACAGCTGTGGAAGAACCGAGCTGACCTGG 880  
 QY 241 GAGTACTGTGATGTGCCCTCTCTGCTCCACCTGCGGCTTGAGACAGTACAGCCGCTCAG 300

DB 881 GAGTACTGTGATGTGCCCTCTCTGCTCCACCTGCGGCTTGAGACAGTACAGCCGCTCAG 940  
 QY 301 TTTGCGCATCAAAGGAGGGCTCTTCGCCGACATCGCTCCCACTCCCTGCGAGCTGCCATC 360  
 DB 941 TTTGCGCATCAAAGGAGGGCTCTTCGCCGACATCGCTCCCACTCCCTGCGAGCTGCCATC 1000  
 QY 361 TTTGCCAAGCACAGGAGGTGCGCCGAGAGCGGTTCTGTGCGGGGGCATACTCATCAGC 420  
 DB 1001 TTTGCCAAGCACAGGAGGTGCGCCGAGAGCGGTTCTGTGCGGGGGCATACTCATCAGC 1060  
 QY 421 TCCTGCTGGATTCTCTCTGCGGCGCCACTGCTTCAGGAGAGGTTTCGCCCCCACCACCTG 480  
 DB 1061 TCCTGCTGGATTCTCTCTGCGGCGCCACTGCTTCAGGAGAGGTTTCGCCCCCACCACCTG 1120  
 QY 481 ACGGTGATCTTGGGCGAGAACATACCGGTGTCCTCGCGAGGAGGAGCAAAATTTGAA 540  
 DB 1121 ACGGTGATCTTGGGCGAGAACATACCGGTGTCCTCGCGAGGAGGAGCAAAATTTGAA 1180  
 QY 541 GTCGAAAAATACATTTGTCATAGGAATTCGATGATGACACTTACGACAATGACATTCGCG 600  
 DB 1181 GTCGAAAAATACATTTGTCATAGGAATTCGATGATGACACTTACGACAATGACATTCGCG 1240  
 QY 601 CTGCTGCACTGAAATCGGATTTCGCTGCTGTCGCCAGGAGAGAGGTTGTCGCACT 660  
 DB 1241 CTGCTGCACTGAAATCGGATTTCGCTGCTGTCGCCAGGAGAGAGGTTGTCGCACT 1300  
 QY 661 GTGTGCTTCCCGCGGCGAGCTCGACGTGCGGACTGGAGCGAGTGTGAGCTCTCCGCGC 720  
 DB 1301 GTGTGCTTCCCGCGGCGAGCTCGACGTGCGGACTGGAGCGAGTGTGAGCTCTCCGCGC 1360  
 QY 721 TACGCGAAGCATGAGGCTTTGTCTCTTTCTATTTCGAGCGGCTGAAAGGAGGCTCATGTC 780  
 DB 1361 TACGCGAAGCATGAGGCTTTGTCTCTTTCTATTTCGAGCGGCTGAAAGGAGGCTCATGTC 1420  
 QY 781 AGACTGTACCATCCAGCGCTGCACATCAACAATTTACTTTACAGAACAGTACCGAC 840  
 DB 1421 AGACTGTACCATCCAGCGCTGCACATCAACAATTTACTTTAAACAGAACAGTACCGAC 1480  
 QY 841 AACATGCTGTGTGCTGAGACACTCGGAGCGGCGGCCCCAGGCAAACTTGCACGAGGCC 900  
 DB 1481 AACATGCTGTGTGCTGAGACACTCGGAGCGGCGGCCCCAGGCAAACTTGCACGAGGCC 1540  
 QY 901 TGCCAGGCGGATTCGGGAGGCGCCCTGCTGTGTCTGAAACGATGGCCGCGATGACTTTGTTG 960  
 DB 1541 TGCCAGGCGGATTCGGGAGGCGCCCTGCTGTGTCTGAAACGATGGCCGCGATGACTTTGTTG 1600  
 QY 961 GGCATCATCAGCTGGGCGCTGGGCTGTGACAGAGGATGTCCCGGGTGTGTACACAAAG 1020  
 DB 1601 GGCATCATCAGCTGGGCGCTGGGCTGTGACAGAGGATGTCCCGGGTGTGTACACAAAG 1660  
 QY 1021 GTTACCAACTACCTAGACTGGATTTCGTGACAACTGCGCCGTGA 1065  
 DB 1661 GTTACCAACTACCTAGACTGGATTTCGTGACAACTGCGCCGTGA 1705  
 RESULT 10  
 AAN91119  
 ID AAN91119 standard; DNA; 2100 BP.  
 XX  
 AC AAN91119;  
 XX  
 AC  
 DT 25-MAR-2003 (updated)  
 DT 03-OCT-2002 (updated)  
 DT 18-JUN-1990 (first entry)  
 XX  
 DE Sequence of native tPA in plasmid pST112.  
 XX  
 KW Tissue plasminogen activator; tPA; thrombolytic agent;  
 KW plasminogen; vascular diseases.  
 XX  
 OS Synthetic.  
 XX

FH Key Location/Qualifiers  
 FT CDS 25..1710  
 FT /\*tag= a  
 XX EP302456-A.  
 XX  
 XX PD 08-FEB-1989.  
 XX  
 XX PF 02-AUG-1988; 88EP-0112569.  
 XX  
 XX PR 03-AUG-1987; 87GB-0018298.  
 XX PR 26-OCT-1987; 87GB-0025052.  
 XX PR 13-NOV-1987; 87GB-0026683.  
 XX  
 XX PA (FUJI ) FUJISAWA PHARM CO LTD.  
 XX  
 XX PI Niwa M, Saito Y, Sasaki H, Hayashi M, Notani J, Kobayaashi M;  
 XX WPI; 1989-040625/06.  
 XX DR P-PSDB; AAP94406.  
 XX  
 XX PT New tissue plasminogen activator -  
 PT comprising finger and growth factor domains lacking tPA for  
 PT longer half-life and stronger thrombolytic activity.  
 XX  
 XX PS Example 29; Fig 21; 68pp; English.  
 XX  
 XX CC cDNA sequence of native tPA gene is excised from plasmid pSt112, and  
 CC digested with BamHI and SalI to form plasmid pSt118.  
 CC (Updated on 03-OCT-2002 to add missing OS field.)  
 CC (Updated on 25-MAR-2003 to correct PA field.)  
 XX  
 XX SQ Sequence 2100 BP; 498 A; 595 C; 597 G; 410 T; 0 other;  
 Query Match 100.0%; Score 1065; DB 10; Length 2100;  
 Best Local Similarity 100.0%; Pred. No. 1.1e-236;  
 Matches 1065; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCTGAGGAAACAGTACTGCTTCTGGGAATGGTCCAGCTACCGTGGCAGCAGC 60  
 DB TCTGAGGAAACAGTACTGCTTCTGGGAATGGTCCAGCTACCGTGGCAGCAGC 708  
 QY 61 CTCACCGAGTGGGTCCTCTGCTCCCTCCGCGAATTCATGATCCCTATAGGCAAGTT 120  
 DB CTCACCGAGTGGGTCCTCTGCTCCCTCCGCGAATTCATGATCCCTATAGGCAAGTT 768  
 QY 121 TACACAGACAGAACCCAGTGGCCAGGACCTGGGCTGGGCAACATAATTAATCTGCCGG 180  
 DB TACACAGACAGAACCCAGTGGCCAGGACCTGGGCTGGGCAACATAATTAATCTGCCGG 828  
 QY 181 AATCTGATGGGATGCCAAGCCCTGCTGCTCCACCTGGGCTGAGACAGTACAGCCTCAG 240  
 DB AATCTGATGGGATGCCAAGCCCTGCTGCTCCACCTGGGCTGAGACAGTACAGCCTCAG 888  
 QY 241 GAGTACTGTGATGTCCTCTGCTCCACCTGGGCTGAGACAGTACAGCAGCCTCAG 300  
 DB GAGTACTGTGATGTCCTCTGCTCCACCTGGGCTGAGACAGTACAGCAGCCTCAG 948  
 QY 301 TTTGCGCATAAAGAGGGGCTTTGCGCGCCACATGCGCTCCACCTGGCAGGCTGCCATC 360  
 DB TTTGCGCATAAAGAGGGGCTTTGCGCGCCACATGCGCTCCACCTGGCAGGCTGCCATC 1008  
 QY 361 TTTGCGAAGCAGAGGAGTCCCGGAGAGCGGTTCTGTCGGGGGCATCTCATCAGC 420  
 DB TTTGCGAAGCAGAGGAGTCCCGGAGAGCGGTTCTGTCGGGGGCATCTCATCAGC 1068  
 QY 421 TCCTGCTGATTTCTCTGCGCGCCACCTGCTTCCAGGAGAGGTTTCCGCCCCACACCTG 480  
 DB TCCTGCTGATTTCTCTGCGCGCCACCTGCTTCCAGGAGAGGTTTCCGCCCCACACCTG 1128  
 QY 481 ACGGTGATCTTGGGCAAGACATACCGGGTGTCTCCCTGGCGAGGAGCAGAAATTTGAA 540  
 DB ACGGTGATCTTGGGCAAGACATACCGGGTGTCTCCCTGGCGAGGAGCAGAAATTTGAA 1188

QY 541 GTCGAAAAATACATTGTCCATTAAGGAATTCGATGATGACACTTACGACAAATGACATTGCG 600  
 DB GTCGAAAAATACATTGTCCATTAAGGAATTCGATGATGACACTTACGACAAATGACATTGCG 1248  
 QY 601 CTGCTGCAGCTGAAATCGGATTCGTCCCGTGTGCCAGGAGAGCAGCGTGGTCCGCACT 660  
 DB CTGCTGCAGCTGAAATCGGATTCGTCCCGTGTGCCAGGAGAGCAGCGTGGTCCGCACT 1308  
 QY 661 GTGTGCTTCCCGCGGAGACCTGACAGTCCGCGACTGGACGAGTGTGAGCTTCCGGC 720  
 DB GTGTGCTTCCCGCGGAGACCTGACAGTCCGCGACTGGACGAGTGTGAGCTTCCGGC 1368  
 QY 721 TACGGCAAGCATGAGGCTTGTCTCTTCTATTTCGGAGCGGCTGAAGAGGCTCATGTC 780  
 DB TACGGCAAGCATGAGGCTTGTCTCTTCTATTTCGGAGCGGCTGAAGAGGCTCATGTC 1428  
 QY 781 AGACTGTACCCATCCAGCCGCTGCACATCAACAATTTACTTAAACAGAACAGTCAACGAC 840  
 DB AGACTGTACCCATCCAGCCGCTGCACATCAACAATTTACTTAAACAGAACAGTCAACGAC 1488  
 QY 841 AACATGCTGTGCTGAGACACTCGGAGCGGCGGCCAGGCAAACTTGCACGACGCC 900  
 DB AACATGCTGTGCTGAGACACTCGGAGCGGCGGCCAGGCAAACTTGCACGACGCC 1548  
 QY 901 TGCCAGGCGGATTCGGGAGGCCCTCTGCTGCTGAACGATGCGCCGCTGCTTGGTG 960  
 DB TGCCAGGCGGATTCGGGAGGCCCTCTGCTGCTGAACGATGCGCCGCTGCTTGGTG 1608  
 QY 961 GGCAATCATAGCTGGGCGCTGGGCTGTGGACAGAGGATGTCCGGGTGTGTACACAAAG 1020  
 DB GGCAATCATAGCTGGGCGCTGGGCTGTGGACAGAGGATGTCCGGGTGTGTACACAAAG 1668  
 QY 1021 GTTACCAACTACCTAGACTCGATTCGTGACACATGTCGACCGTGA 1065  
 DB GTTACCAACTACCTAGACTCGATTCGTGACACATGTCGACCGTGA 1713  
 RESULT 11  
 AAQ05532  
 ID AAQ05532 standard; DNA; 2100 BP.  
 XX AC AAQ05532;  
 XX AC  
 XX 25-MAR-2003 (updated)  
 DT 11-DEC-1990 (first entry)  
 XX  
 DE Plasmid pSt112 encoding novel N-terminal for tissue plasminogen  
 DE activator (tPA).  
 XX  
 KW Fibrin; clotting; thrombolytic; vascular disease; stroke;  
 KW myocardial infarction; heart attack; pulmonary embolism; ds;  
 XX Homo sapiens.  
 OS  
 XX  
 XX Key Location/Qualifiers  
 FT CDS 25..1710  
 FT /\*tag= a  
 FT mat\_peptide 130..1710  
 FT /\*tag= b  
 XX  
 XX EP379890-A.  
 XX  
 XX PD 01-AUG-1990.  
 XX  
 XX PF 10-JAN-1990; 90BP-0100457.  
 XX  
 XX PR 23-JAN-1989; 89GB-0001422.  
 XX  
 XX PA (FUJI ) FUJISAWA PHARM CO LTD.  
 XX  
 XX NIWA M, SATOH S, SUZUKI S, OTSUKA K, KUSUNOKI C;  
 XX

DR WPI: 1990-232757/31.  
XX P-PSDB; AAR06237.  
PT New tissue plasminogen activator - having N-terminal peptide of  
PT plasminogen linked to tissue plasminogen activator for increased  
PT stability in vivo.  
XX  
XX  
PS Disclosure; Page ?; ?pp; English.  
CC tPA with N-terminal peptide of plasminogen linked is more stable in  
CC vivo than the native form. It is useful as a thrombolytic agent in  
CC the treatment of vascular diseases eg myocardial infarction,  
CC pulmonary embolism etc.  
CC (Updated on 25-MAR-2003 to correct PA field.)  
XX  
XX  
SQ Sequence 2100 BP; 498 A; 595 C; 597 G; 410 T; 0 other;  
Query Match 100.0%; Score 1065; DB 11; Length 2100;  
Best Local Similarity 100.0%; Pred. No. 1.1e-236;  
Matches 1065; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
1 TCTGAGGAAACAGTGACTGCTACTTTGGGAATGGGTGAGCTACCGTGGCAGCAGCAGC 60  
649 TCTGAGGAAACAGTGACTGCTACTTTGGGAATGGGTGAGCTACCGTGGCAGCAGCAGC 708  
61 CTCACGAGTGGGTGCTCTGCTCCCTCCGTGGGAATTCATGATCCTGATAGGCAAGTT 120  
709 CTCACGAGTGGGTGCTCTGCTCCCTCCGTGGGAATTCATGATCCTGATAGGCAAGTT 768  
121 TACACGACACAGAACCCAGTGCCTCCAGCAGCTGGCCCTGGCCAAACATAATTAAGTCCGG 180  
769 TACACGACACAGAACCCAGTGCCTCCAGCAGCTGGCCCTGGCCAAACATAATTAAGTCCGG 828  
181 AATCTGATGGGATGCCAAGCCCTGGTGGCAGCTGCTGAAGAACCGCAGCTGACGTGG 240  
829 AATCTGATGGGATGCCAAGCCCTGGTGGCAGCTGCTGAAGAACCGCAGCTGACGTGG 888  
241 GAGTACTGTGATGTCCTCTGCTCCACCTGGCCCTGAGACAGTACAGCCAGCTCAG 300  
889 GAGTACTGTGATGTCCTCTGCTCCACCTGGCCCTGAGACAGTACAGCCAGCTCAG 948  
301 TTTGCAATCAAGAGGAGCTCTTCCGACATCCGCTCCACCCCTGGCAGGCTGCCATC 360  
949 TTTGCAATCAAGAGGAGCTCTTCCGACATCCGCTCCACCCCTGGCAGGCTGCCATC 1008  
361 TTTGCAATCAAGAGGAGCTCTTCCGACATCCGCTCCACCCCTGGCAGGCTGCCATC 420  
1009 TTTGCAATCAAGAGGAGCTCTTCCGACATCCGCTCCACCCCTGGCAGGCTGCCATC 1068  
421 TCCTGCTGGATCTCTTCCGACATCCGCTCCACCCCTGGCAGGAGGTTTCCGCCCAACCTG 480  
1069 TCCTGCTGGATCTCTTCCGACATCCGCTCCACCCCTGGCAGGAGGTTTCCGCCCAACCTG 1128  
481 ACGGTGATCTTGGCAGAACATACCGGTGCTCCCTGGCAGGAGGAGCAGAAATTGAA 540  
1129 ACGGTGATCTTGGCAGAACATACCGGTGCTCCCTGGCAGGAGGAGCAGAAATTGAA 1188  
541 CTCGAAATAATATGTCATAAGGAATTCGATGATGACCTTACGACAAATGACATTGCG 600  
1189 CTCGAAATAATATGTCATAAGGAATTCGATGATGACCTTACGACAAATGACATTGCG 1248  
601 CTCGTGACGTGAATTCGGATTCGCTCCGCTGTGCCCAGGAGAGCAGCTGGTCCGCACT 660  
1249 CTCGTGACGTGAATTCGGATTCGCTCCGCTGTGCCCAGGAGAGCAGCTGGTCCGCACT 1308  
661 GTGTGCTTCCCGCGGAGCTGAGCTGCGGAGCTGAGAGAGGAGCTGAGCTTCCGGC 720  
1309 GTGTGCTTCCCGCGGAGCTGAGCTGCGGAGCTGAGAGAGGAGCTGAGCTTCCGGC 1368  
721 TACGCAAGCATGAGGCTTGTCTCTTCTTATTCGGAGCGCTGAAGAGGCTCATGTC 780  
1369 TACGCAAGCATGAGGCTTGTCTCTTCTTATTCGGAGCGCTGAAGAGGCTCATGTC 1428

QY 781 AGACTGTACCATCCAGCCGCTGCACATCAACAATTTACTTAACAGAACAGTCAACGAC 840  
DB 1429 AGACTGTACCATCCAGCCGCTGCACATCAACAATTTACTTAACAGAACAGTCAACGAC 1488  
QY 841 AACATGCTGTGTCTGGAGACACTCGGAGCGGCGGCCCGGCAAACTTGCACGACGCC 900  
DB 1489 AACATGCTGTGTCTGGAGACACTCGGAGCGGCGGCCCGGCAAACTTGCACGACGCC 1548  
QY 901 TSCCAGGCGGATTCGGGAGGCCCTCTGCTGTGTCTGAAACGATGGCCGATGACTTTGGTG 960  
DB 1549 TGCCAGGCGGATTCGGGAGGCCCTCTGCTGTGTCTGAAACGATGGCCGATGACTTTGGTG 1608  
QY 961 GGCATCATCAGCTGGGCGCTGGGCTGTGGACAGAGGATGTCGGGTGTGTACACAAG 1020  
DB 1609 GGCATCATCAGCTGGGCGCTGGGCTGTGGACAGAGGATGTCGGGTGTGTACACAAG 1668  
QY 1021 GTTACCAACTACCTAGACTGGATTCTGACAAACATGCGACCGTGA 1065  
DB 1669 GTTACCAACTACCTAGACTGGATTCTGACAAACATGCGACCGTGA 1713  
RESULT 12  
AAQ87370  
ID AAQ87370 standard; DNA; 2162 BP.  
XX  
AC AAQ87370;  
XX AC  
XX 25-MAR-2003 (updated)  
DT 19-SEP-1995 (first entry)  
XX  
XX Human tPA clone.  
XX  
KW Intron; recombination; combinatorial gene; trans-splicing;  
KW gene therapy; polymerase chain reaction; PCR; primer; amplification;  
KW tissue plasminogen activator; tPA; plasmid TPA-KS; thrombolytic;  
KW ds.  
XX  
XX Synthetic.  
XX  
XX Key Location/Qualifiers  
FT sig\_peptide /tag= a  
FT /note= "signal sequence and finger-like domain"  
FT CDS /tag= b  
FT /product= EGF-like domain  
FT CDS /tag= c  
FT /product= Kringle-1 domain  
FT CDS /tag= d  
FT /product= Kringle-2 domain  
FT CDS /tag= e  
FT /product= catalytic domain  
XX  
XX W09507351-A1.  
XX  
XX 16-MAR-1995.  
XX  
XX 12-SEP-1994; 94WO-US10146.  
XX  
XX 10-SEP-1993; 93US-0119512.  
XX  
XX (HARD ) HARVARD COLLEGE.  
XX  
XX Jarrell KA;  
XX  
XX WPI; 1995-123425/16.  
XX  
XX  
XX New intron-mediated recombinant techniques - used for the  
XX generation and selection of novel genes and gene prods. for use  
XX in therapy  
PT

XX  
PS  
XX  
XX  
CC A cDNA clone of human tissue plasminogen activator (tPA) was  
CC amplified by PCR using the primers given in AA087368-69. The  
CC amplified tPA DNA (AA087370) was ligated into vector KS+ to  
CC obtain plasmid TPA-KS+. The construct was used in combinatorial  
CC methods involving RNA splicing-mediated shuffling of tPA domains  
CC in plasmid pINVI (AA087347) to generate novel tPAs having  
CC improved thrombolytic properties.  
CC (Updated on 25-MAR-2003 to correct PN field.)  
XX  
SQ Sequence 2162 BP; 520 A; 606 C; 619 G; 417 T; 0 other;  
  
Query Match 100.08; Score 1065; DB 16; Length 2162;  
Best Local Similarity 100.08; Pred. No. 1.1e-236;  
Matches 1065; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 TCTGAGGAAACAGTGAAGTCTTCTGGGAATGGTCCAGCTACCGTGGCAGCAGCAGC 60  
DB TCTGAGGAAACAGTGAAGTCTTCTGGGAATGGTCCAGCTACCGTGGCAGCAGCAGC 765  
  
QY 61 CTCACCGAGTGGGGTCTCTGCTCCCGTGGGAATTCATGATCCTGATAGGCAAGTT 120  
DB CTCACCGAGTGGGGTCTCTGCTCCCGTGGGAATTCATGATCCTGATAGGCAAGTT 825  
  
QY 121 TACACGACACAGAACCCAGTCCCGAGCAGCTGGGCTGGGCAACATAATCTGCGG 180  
DB TACACGACACAGAACCCAGTCCCGAGCAGCTGGGCTGGGCAACATAATCTGCGG 885  
  
QY 181 AATCTGATGGGATGCCAAGCCCTGGTGGCCAGCTGGTGAAGAACCGCAGCTGACGTGG 240  
DB AATCTGATGGGATGCCAAGCCCTGGTGGCCAGCTGGTGAAGAACCGCAGCTGACGTGG 945  
  
QY 241 GAGTACTGTGATGTCCTCTGCTCCACCTGGCCCTTGAGACATACAGCAAGCTTCAG 300  
DB GAGTACTGTGATGTCCTCTGCTCCACCTGGCCCTTGAGACATACAGCAAGCTTCAG 1005  
  
QY 301 TTTCCGATCAAGAGGAGGCTCTTCCGCGACATCGCTCCACCCCTGGCAGGCTGCCATC 360  
DB TTTCCGATCAAGAGGAGGCTCTTCCGCGACATCGCTCCACCCCTGGCAGGCTGCCATC 1065  
  
QY 361 TTTGCCAAGCAGAGAGTCCGCCCGAGGAGGCTTCTGTGCGGGGATCTACTCATCAGC 420  
DB TTTGCCAAGCAGAGAGTCCGCCCGAGGAGGCTTCTGTGCGGGGATCTACTCATCAGC 1125  
  
QY 421 TCCTGCTGGATTTCTCTGCGGCCCACTGCTTCCAGGAGAGTTTCCGCCCAACCACTG 480  
DB TCCTGCTGGATTTCTCTGCGGCCCACTGCTTCCAGGAGAGTTTCCGCCCAACCACTG 1185  
  
QY 481 ACGGTGATCTGGGCGAGACATACCGGCTGGTCCCTGGCGAGGAGCAGAAATTTGAA 540  
DB ACGGTGATCTGGGCGAGACATACCGGCTGGTCCCTGGCGAGGAGCAGAAATTTGAA 1245  
  
QY 541 GTCGAAAATATATGTGCATAAGGAATTCGATGATGACATTCAGCAATGACATTCGCG 600  
DB GTCGAAAATATATGTGCATAAGGAATTCGATGATGACATTCAGCAATGACATTCGCG 1305  
  
QY 601 CTGCTGCAAGTGAATTCGATTCGCTCCGCTGTGCTCCAGGAGAGCAGCTGGTCCGCACT 660  
DB CTGCTGCAAGTGAATTCGATTCGCTCCGCTGTGCTCCAGGAGAGCAGCTGGTCCGCACT 1365  
  
QY 661 GTGTGCTCTCCCGCGGAGCTGAGCTGGGAGCTGGAGCTGAGCTCTCCGGC 720  
DB GTGTGCTCTCCCGCGGAGCTGAGCTGGGAGCTGGAGCTGAGCTCTCCGGC 1425  
  
QY 721 TACGCAAGCATGAGGCTTGTCTCTTCTATTCGAGCGGCTGAAGAGGCTCATGCTC 780  
DB TACGCAAGCATGAGGCTTGTCTCTTCTATTCGAGCGGCTGAAGAGGCTCATGCTC 1485  
  
QY 781 AGACTGTACCCATCCAGCGGCTGCAATCAACAATTTACTTAAACAGAACAGTCAACGAC 840  
DB AGACTGTACCCATCCAGCGGCTGCAATCAACAATTTACTTAAACAGAACAGTCAACGAC 1545

QY 841 AACATGCTGTGTGCTGGAGACACTCGGAGCGGGGCCCCAGGCAAACTTGCACGAGCC 900  
DB AACATGCTGTGTGCTGGAGACACTCGGAGCGGGGCCCCAGGCAAACTTGCACGAGCC 1605  
  
QY 901 TGGCAGGCGGATTCGGAGGCCCCCTGCTGTGTGTTGAACGATGGCCCATGACTTTGGTG 960  
DB TGGCAGGCGGATTCGGAGGCCCCCTGCTGTGTGTTGAACGATGGCCCATGACTTTGGTG 1665  
  
QY 961 GGCATCATCAGCTGGGCGCTGGGCTGTGGACAGAGATGTCCCGGCTGTGTACACAAAG 1020  
DB GGCATCATCAGCTGGGCGCTGGGCTGTGGACAGAGATGTCCCGGCTGTGTACACAAAG 1725  
  
QY 1021 GTTACCAACTACCTAGACTGGATTTCGTGACAAACATGCGACCGTGA 1065  
DB GTTACCAACTACCTAGACTGGATTTCGTGACAAACATGCGACCGTGA 1770  
  
RESULT 13  
AAV37294  
ID AAV37294 standard; DNA; 2162 BP.  
XX  
AC AAV37294;  
XX  
DT 10-SEP-1998 (first entry)  
XX  
DE Human tissue plasminogen activator gene sequence.  
XX  
KW Plasmid pINVI; reverse-splicing intron; group II intron;  
KW exon binding site; domain V motif; branch site acceptor;  
KW nucleophilic group; transesterification; phosphodiester bond;  
KW autocatalytic Y-branched intron; reverse splicing reaction; ds.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT misc\_feature 82..334  
FT FT /\*tag= a  
FT FT /note= "signal sequence and finger-like domain".  
FT FT misc\_feature 335..447  
FT FT /\*tag= b  
FT FT /note= "EGF-like domain"  
FT FT misc\_feature 448..714  
FT FT /\*tag= c  
FT FT /note= "Kringle-1 domain"  
FT FT misc\_feature 715..972  
FT FT /\*tag= d  
FT FT /note= "Kringle-2 domain"  
FT FT misc\_feature 973..2162  
FT FT /\*tag= e  
FT FT /note= "catalytic domain"  
XX  
PN US5780272-A.  
XX  
PD 14-JUL-1998.  
XX  
PP 07-JUN-1995; 95US-0488015.  
XX  
PR 10-SEP-1993; 93US-0119512.  
XX  
PA (HARD ) HARVARD COLLEGE.  
XX  
PI Jarrell KA;  
XX  
DR WPI; 1998-413060/35.  
XX  
PT Reverse splicing construct containing fragments of autocatalytic  
PT introns - able to cleave and ligate discontinuous nucleic acid for  
PT generating new genes and e.g. ribozymes, libraries of enzymes and  
PT antibodies  
XX  
PS Example 4; Columns 53-56; 56pp; English.  
XX

The present sequence represents the human tissue plasminogen activator gene. It was used to construct plasmid TPA-KS+, which is used in the course of the invention. The specification describes a purified reverse-splicing intron which comprises a segment comprising a 5'-part of a group II intron, including an exon binding site not naturally present in the intron and a second segment comprising a 3'-part of a group II intron, including a domain V motif, a branch site acceptor, and a nucleophilic group for transesterifying a phosphodiester bond of an RNA. Together the two segments form an autocatalytic Y-branched intron which catalyses integration of at least the first segment into substrate RNA by a reverse splicing reaction. The reverse-splicing introns are used, by specific cleavage and ligation of discontinuous nucleic acid, to generate new genes and gene products, e.g. ribozymes (for use in gene therapy or as reagents in DNA manipulation, e.g. replacements for restriction enzymes) or immunologically active or signal-transducing proteins such as antibody and enzyme libraries.

SQ Sequence 2162 BP; 520 A; 606 C; 619 G; 417 T; 0 other;

Query Match 100.0%; Score 1065; DB 19; Length 2162;

Best Local Similarity 100.0%; Pred. No. 1.1e-236;

Matches 1065; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCTGAGGAAACAGTACTGCTACTTTGGGAATGGGTGAGCTACCTACCGTGCGACGACAGC 60  
 DB 706 TCTGAGGAAACAGTACTGCTACTTTGGGAATGGGTGAGCTACCTACCGTGCGACGACAGC 765  
 QY 61 CTCACCGAGTGGGTGCTCTCCCTCCCTGCGTGAATTCATGATCCTGATAGGCAAGTT 120  
 DB 766 CTCACCGAGTGGGTGCTCTCCCTCCCTGCGTGAATTCATGATCCTGATAGGCAAGTT 825  
 QY 121 TACACGACAGAACCCAGTCCCGAGGCACTGGGCTGGGCAACATAATTAATCTGCGCG 180  
 DB 826 TACACGACAGAACCCAGTCCCGAGGCACTGGGCTGGGCAACATAATTAATCTGCGCG 885  
 QY 181 ATCTCTGATGGGATGCCAGCCCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240  
 DB 886 ATCTCTGATGGGATGCCAGCCCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 945  
 QY 241 GAGTACTGTGATGCTCCCTCTCCCTCCCTGCGTGAATTCATGATCCTGATAGGCAAGTT 300  
 DB 946 GAGTACTGTGATGCTCCCTCTCCCTCCCTGCGTGAATTCATGATCCTGATAGGCAAGTT 1005  
 QY 301 TTTGCGATCAAGAGGAGGCTCTTCGCGACATCGCTCCCGCTGGCAGGCTGCCATC 360  
 DB 1006 TTTGCGATCAAGAGGAGGCTCTTCGCGACATCGCTCCCGCTGGCAGGCTGCCATC 1065  
 QY 361 TTTGCGACAGCAGGAGTGGCGCGGAGAGCGGTTCTGTGCGGGGATATCTCATCAGC 420  
 DB 1066 TTTGCGACAGCAGGAGTGGCGCGGAGAGCGGTTCTGTGCGGGGATATCTCATCAGC 1125  
 QY 421 TCCTGCTGATTTCTCTGCGCGCCACTGCTTCCAGGAGAGGTTTCCCGCCCAACACCTG 480  
 DB 1126 TCCTGCTGATTTCTCTGCGCGCCACTGCTTCCAGGAGAGGTTTCCCGCCCAACACCTG 1185  
 QY 481 ACGGTGATCTGGGCGAGAACATACCGGGTGTCTTCGCGAGGAGGAGCAAAATTTGAA 540  
 DB 1186 ACGGTGATCTGGGCGAGAACATACCGGGTGTCTTCGCGAGGAGGAGCAAAATTTGAA 1245  
 QY 541 GTCGAAATAATGATTTGTCATGAAGAAATTCGATGATGACACTTACGACAAATGACATTCG 600  
 DB 1246 GTCGAAATAATGATTTGTCATGAAGAAATTCGATGATGACACTTACGACAAATGACATTCG 1305  
 QY 601 CTGCTGACGCTGAATTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660  
 DB 1306 CTGCTGACGCTGAATTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1365  
 QY 661 GTGTGCTTCCCGCGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720  
 DB 1366 GTGTGCTTCCCGCGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1425  
 QY 721 TACGCAAGCATGAGGCTTGTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 780

DB 1426 TACGCAAGCATGAGGCTTGTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1485  
 QY 781 AGACTGTACCATCCAGCCGCTGCACATCAACAATTTACTTAACAGAACAGTCAACCGAC 840  
 DB 1486 AGACTGTACCATCCAGCCGCTGCACATCAACAATTTACTTAACAGAACAGTCAACCGAC 1545  
 QY 841 AACATGCTGTGTCTGGAGACACTCGGAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 900  
 DB 1546 AACATGCTGTGTCTGGAGACACTCGGAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1605  
 QY 901 TCCAGGCGGATTCGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 960  
 DB 1606 TCCAGGCGGATTCGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1665  
 QY 961 GGCATCATCAGTGGGCGCTGGGCTGTGACAGAGGATGTCGCGGGTGTGTACACAAG 1020  
 DB 1666 GGCATCATCAGTGGGCGCTGGGCTGTGACAGAGGATGTCGCGGGTGTGTACACAAG 1725  
 QY 1021 GTTACCAACTACCTAGACTGGATTCTGTGACAACTGCGACCGTGA 1065  
 DB 1726 GTTACCAACTACCTAGACTGGATTCTGTGACAACTGCGACCGTGA 1770

# RESULT 14

AA60659

ID AA60659 standard; DNA; 2547 BP.

XX AA60659;

XX 25-MAR-2003 (updated)

DT 22-JUL-1991 (first entry)

XX Sequence encoding human pre-tissue plasminogen activator (pre-t-PA).

XX Plasminogen conversion; vascular disease therapy; ss.

XX Homo sapiens.

XX Key

XX CDS

XX mat\_peptide

XX GB2173804-A.

XX 22-OCT-1986.

XX 21-APR-1986; 86GB-0609683.

XX 22-APR-1985; 85US-0725468.

XX 01-APR-1986; 86US-0846697.

XX 22-APR-1985; 85US-0725468.

XX 01-APR-1986; 86US-0846697.

XX (GETH ) GENENTECH INC.

XX Heyneker HL, Vehar GA;

XX WPI; 1986-280715/43.

XX P-PSDB; AAP60790.

XX New mutant forms of human tissue plasminogen activator- having

XX higher specific activity and resistance to conversion to two chain

XX form

XX Example; Fig 2; 34pp; English.

XX The patentors claim a novel recombinant human t-PA which is

XX resistant to specific enzymatic cleavage because it is stabilised by

XX site-directed mutagenesis at a 2-chain cleavage site. Partic. the

XX natural Arg at position 275 is replaced by Gly or Glu, or Ile at

CC position 276 is replaced (275 and 276 refer to the posn. of the AAs  
CC in the mature protein; i.e. AAs 310 and 311 of AAP60790). Also new are  
CC (1) DNA sequences; (2) expression vectors; and (3) microorganisms  
CC and cell cultures transformed with these vectors.  
CC (Updated on 25-MAR-2003 to correct PR field.)  
XX

Sequence 2547 BP; 627 A; 688 C; 691 G; 541 T; 0 other;

Query Match 100.0%; Score 1065; DB 7; Length 2547;  
Best Local Similarity 100.0%; Pred. No. 1.1e-236; Indels 0; Gaps 0;  
Matches 1065; Conservative 0; Mismatches 0;

Qy 1 TCTGAGGAAACAGTACTGCTACTTTTGGGAATGGGTGAGCTACCGTGGCAGCAGCAGC 60  
Db TCTGAGGAAACAGTACTGCTACTTTTGGGAATGGGTGAGCTACCGTGGCAGCAGCAGC 770  
Qy 61 CTCACCGAGTGGGTGCTCTCTGCTCCGCTGGGAATTCATGATCTCTATAGGCAAGTTT 120  
Db CTCACCGAGTGGGTGCTCTCTGCTCCGCTGGGAATTCATGATCTCTATAGGCAAGTTT 830  
Qy 121 TACACGACACAGAACCCAGTGCAGCAGCAGTGGGCTGGGCAACATAATTAATCTGCGG 180  
Db TACACGACACAGAACCCAGTGCAGCAGCAGTGGGCTGGGCAACATAATTAATCTGCGG 890  
Qy 181 AATCTGTATGGGATGCCAAGCCCTGTGTCACAGTCTGCTGAAGAACCGCAGGCTGACGTGG 240  
Db AATCTGTATGGGATGCCAAGCCCTGTGTCACAGTCTGCTGAAGAACCGCAGGCTGACGTGG 950  
Qy 241 GAGTACTGTGATGTCCTCTCTGCTCCACTGCGGCTTGAGACAGTACAGCAGCCTCAG 300  
Db GAGTACTGTGATGTCCTCTCTGCTCCACTGCGGCTTGAGACAGTACAGCAGCCTCAG 1010  
Qy 301 TTTTGCATCAAGAGGAGGCTCTTCCGACATCGCTCCACCCCTGCGAGGCTGCCATC 360  
Db TTTTGCATCAAGAGGAGGCTCTTCCGACATCGCTCCACCCCTGCGAGGCTGCCATC 1070  
Qy 361 TTTTGCACAGCAGAGGTCGCGCGGAGAGGGTTCTGTGCGGGGATACTCATCAGC 420  
Db TTTTGCACAGCAGAGGTCGCGCGGAGAGGGTTCTGTGCGGGGATACTCATCAGC 1130  
Qy 421 TCCTGCTGGATTTCTCTGCGGCCCACTGCTTCCAGGAGAGGTTTCCGCCCAACACCTG 480  
Db TCCTGCTGGATTTCTCTGCGGCCCACTGCTTCCAGGAGAGGTTTCCGCCCAACACCTG 1130  
Qy 481 ACGGTGATCTGGGACAGACATACCGGTGCTCCTGCGGAGGAGGACGAAATTTGAA 540  
Db ACGGTGATCTGGGACAGACATACCGGTGCTCCTGCGGAGGAGGAGGACGAAATTTGAA 1250  
Qy 541 GTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAAATGACATTGCG 600  
Db GTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAAATGACATTGCG 1310  
Qy 601 CTGCTGACGCTGAATTCGGATTCGCTCGCTGTGCGCCAGAGAGCAGGTCGGTCCGCACT 660  
Db CTGCTGACGCTGAATTCGGATTCGCTCGCTGTGCGCCAGAGAGCAGGTCGGTCCGCACT 1370  
Qy 661 GTGTCCTTCCCGCGGACCTGAGTCCGCGTGGGAGTGGAGTGGAGTGGAGTGGAGTGGAGT 720  
Db GTGTCCTTCCCGCGGACCTGAGTCCGCGTGGGAGTGGAGTGGAGTGGAGTGGAGTGGAGT 1430  
Qy 721 TACGCAAGCATGAGGCTTGTCTCTCTTTATTCGAGCGGCTGAAGAGGCTCATGTC 780  
Db TACGCAAGCATGAGGCTTGTCTCTCTTTATTCGAGCGGCTGAAGAGGCTCATGTC 1490  
Qy 781 AGACTGTACCCATTCAGCGCTGCAATCAACATTTACTTAAACAGAAAGTACCGAC 840  
Db AGACTGTACCCATTCAGCGCTGCAATCAACATTTACTTAAACAGAAAGTACCGAC 1550  
Qy 841 AACATGCTGTGCTGGAGACACTCGGAGCGCGGCCCGGCAAACTTCCAGCAGGCC 900  
Db AACATGCTGTGCTGGAGACACTCGGAGCGCGGCCCGGCAAACTTCCAGCAGGCC 1610  
Qy 901 TGCCAGGGCGATTGCGGAGGGCCCCCTGGTGTCTCTGAACGATGGCGCATGACTTTTGGT 960

Db 1611 TGCCAGGCGGATTTCGGAGGCCCTCTGTGTCTGAACGATGGCCGATGACTTTGGTG 1670  
Qy 961 GGCATCATCAGCTGGGCGCTGGGCTGGGAGAGATGTCCGGGTGTGTACACAAAG 1020  
Db 1671 GGCATCATCAGCTGGGCGCTGGGCTGGGAGAGATGTCCGGGTGTGTACACAAAG 1730  
Qy 1021 GTTACCAACTACCTAGACTGGATTTCGTGACAAACATGCGACCGTGA 1065  
Db 1731 GTTACCAACTACCTAGACTGGATTTCGTGACAAACATGCGACCGTGA 1775

## RESULT 15

AAN90542

ID AAN90542 standard; cDNA; 2560 BP.

XX AAN90542;

AC AAN90542;

XX 25-MAR-2003 (updated)

DT 05-JUN-1990 (first entry)

XX Plasmid pKG12 contg. DNA encoding human melanoma t-PA.

XX Human tissue plasminogen-activator gene; pKG12; ss.

XX Homo sapiens.

XX Key Location/Qualifiers

FT sig\_peptide /tag= a

FT mat\_peptide /tag= b

FT /product=t-PA

XX BP297066-A.

XX 28-DEC-1988.

XX 14-JUN-1988; 88EP-0850207.

XX 18-JUN-1987; 87SE-0002562.

XX (KABI ) KABIGEN AB.

XX Pohl G, Hansson L, Loewenadler B;

XX WPI; 1989-001503/01.

XX P-PSDB; AAP93716.

XX Modified tissue plasminogen-activator - having domains deleted and amino acid changes to increase biological half-life and reduce inactivation.

XX Disclosure; Fig 1; 18pp; English.

XX The sequence was obtained from cDNA prepd. from mRNA extracted from Bowes melanoma cells. The cDNA was used to construct a gene library which was screened with a partial t-PA cDNA clone. Plasmid pKG12 was isolated and shown to comprise the whole coding region for human t-PA as well as 102bp 5'flanking, 760 bp 3' flanking DNA and a poly A tail.

XX See also AAN91608.

XX (Updated on 25-MAR-2003 to correct PI field.)

XX Sequence 2560 BP; 634 A; 690 C; 689 G; 547 T; 0 other;

Query Match 100.0%; Score 1065; DB 10; Length 2560;

Best Local Similarity 100.0%; Pred. No. 1.1e-236;

Matches 1065; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCTGAGGAAACAGTACTGCTACTTTTGGGAATGGGTGAGCTACCGTGGCAGCAGCAGC 60

Db 726 TCTGAGGAAACAGTACTGCTACTTTTGGGAATGGGTGAGCTACCGTGGCAGCAGCAGC 785

Qy 61 CTCACCGAGTGGGTGCTCTCTGCTCCCGTGGAAATTCATGATCTCTATAGGCAAGTTT 120



Db 786 CTCACCGAGTCGGGTGCGCTTCCTGCGCTCCCGTGGAAATTCATGATCCTGATAGGCAAGGTT 845  
QY 121 TACACAGACAGAACCCAGATGCCAGGCACTGGGCGCTGGCAACATAAATTTACTGCCGG 180  
Db 846 TACACAGACAGAACCCAGATGCCAGGCACTGGGCGCTGGCAACATAAATTTACTGCCGG 905  
QY 181 AATCTGATGGGATGCCAGCCCTGGTGGCCACAGTGTGAAGAACCGCAGGCTGACGTGG 240  
Db 906 AATCTGATGGGATGCCAGCCCTGGTGGCCACAGTGTGAAGAACCGCAGGCTGACGTGG 965  
QY 241 GAGTACTGTGATGTCCTCTCCCTGCTCCACCTGGCGGCTGAGACAGTACAGCCAGCCTCAG 300  
Db 966 GAGTACTGTGATGTCCTCTCCCTGCTCCACCTGGCGGCTGAGACAGTACAGCCAGCCTCAG 1025  
QY 301 TTTCCGATCAAGAGGAGGCTCTTCCCGGACATCGCTCCACCCCTGGCAGGCTGCCATC 360  
Db 1026 TTTCCGATCAAGAGGAGGCTCTTCCCGGACATCGCTCCACCCCTGGCAGGCTGCCATC 1085  
QY 361 TTTGCCAAGCAGAGAGTCCGCCCGGAGAGCGGTTCTGTGGGGGGCATATCATCAGC 420  
Db 1086 TTTGCCAAGCAGAGAGTCCGCCCGGAGAGCGGTTCTGTGGGGGGCATATCATCAGC 1145  
QY 421 TCCTGCTGGATCTCTGCGCGCCACTGCTTCCAGGAGAGGTTTCCGCCCCACCCACTG 480  
Db 1146 TCCTGCTGGATCTCTGCGCGCCACTGCTTCCAGGAGAGGTTTCCGCCCCACCCACTG 1205  
QY 481 ACGGTGATCTTGGGAGAACATACCGGGTGTGCTCCGCGAGGAGGAGCAAAATTTGAA 540  
Db 1206 ACGGTGATCTTGGGAGAACATACCGGGTGTGCTCCGCGAGGAGGAGCAAAATTTGAA 1265  
QY 541 GTCGAAAAATACATTTGTCATAAGGAAATTCGATGATGACACTTACGACAATGACATTGCG 600  
Db 1266 GTCGAAAAATACATTTGTCATAAGGAAATTCGATGATGACACTTACGACAATGACATTGCG 1325  
QY 601 CTGCTGAGAGTGAATTCGGATTCGTCCTGCTGTGCGCCAGGAGAGAGCGTGGTCCGCACT 660  
Db 1326 CTGCTGAGAGTGAATTCGGATTCGTCCTGCTGTGCGCCAGGAGAGAGCGTGGTCCGCACT 1385  
QY 661 GTGTGCTCTTCCCGCGGAGCTGAGCTGCGGAGTGGAGGAGTGGAGTCTCCGGC 720  
Db 1386 GTGTGCTCTTCCCGCGGAGCTGAGCTGCGGAGTGGAGGAGTGGAGTCTCCGGC 1445  
QY 721 TACGGCAAGCATAGGCGCTTGTCTCTTTCTATTTCGAGCGGCTGAAGGAGGCTCATGTC 780  
Db 1446 TACGGCAAGCATAGGCGCTTGTCTCTTTCTATTTCGAGCGGCTGAAGGAGGCTCATGTC 1505  
QY 781 AGACTGTACCCATCCAGCGGCTGCACATCAACAATTTACTTACAGAACAGTACCGAC 840  
Db 1506 AGACTGTACCCATCCAGCGGCTGCACATCAACAATTTACTTACAGAACAGTACCGAC 1565  
QY 841 AACATGCTGTGTGCTGGAGACACTCGGAGCGGCGGCGCCAGGCAAACTTGCACGAGCC 900  
Db 1566 AACATGCTGTGTGCTGGAGACACTCGGAGCGGCGGCGCCAGGCAAACTTGCACGAGCC 1625  
QY 901 TGCCAGGCGGATTCGGGAGGCGCCCTGTGTGTCTGAAAGATGGCCGATGACTTTGGTG 960  
Db 1626 TGCCAGGCGGATTCGGGAGGCGCCCTGTGTGTCTGAAAGATGGCCGATGACTTTGGTG 1685  
QY 961 GGCATCATCAGTGGGGCTGGGCTGTGGACAGAGGATGTCCCGGGTGTGTACACAAAG 1020  
Db 1686 GGCATCATCAGTGGGGCTGGGCTGTGGACAGAGGATGTCCCGGGTGTGTACACAAAG 1745  
QY 1021 GTTACCAACTACCTAGACTGGATTCTGTCACAACATGCGACCGTGA 1065  
Db 1746 GTTACCAACTACCTAGACTGGATTCTGTCACAACATGCGACCGTGA 1790

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 18, 2003, 22:51:23 ; Search time 83.3499 Seconds  
(without alignments)  
5639.757 Million cell updates/sec

Title: US-09-987-455-7

Perfect score: 1065

Sequence: 1 tctgagggaacagtgcgtg.....gtgacaacatgcgcacctga 1065

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA.\*

- 1: /cgn2\_6/ptodata/2/ina/5A\_COMB.seq.\*
- 2: /cgn2\_6/ptodata/2/ina/5B\_COMB.seq.\*
- 3: /cgn2\_6/ptodata/2/ina/6A\_COMB.seq.\*
- 4: /cgn2\_6/ptodata/2/ina/6B\_COMB.seq.\*
- 5: /cgn2\_6/ptodata/2/ina/PCTUS\_COMB.seq.\*
- 6: /cgn2\_6/ptodata/2/ina/backfiles.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1065	100.0	1068	2	US-08-811-949-44
2	1065	100.0	1419	2	US-08-811-949-62
3	1065	100.0	1848	3	US-08-814-412-10
4	1065	100.0	2101	2	US-08-811-949-42
5	1065	100.0	2162	1	US-08-119-512-3
6	1065	100.0	2162	1	US-08-488-015B-3
7	1065	100.0	2162	1	US-08-488-015B-25
8	1063.4	99.8	1170	2	US-08-811-949-64
9	1063.4	99.8	1314	2	US-08-811-949-48
10	1063.4	99.8	1738	6	US-08-811-949-48
11	1063.4	99.8	1955	2	US-08-883-795A-39
12	1063.4	99.8	2457	6	5344773-1
13	1063.4	99.8	7360	1	US-08-286-740-1
14	1063.4	99.8	7360	5	PCT-US95-09576-1
15	1061.8	99.7	1068	2	US-08-811-949-46
16	1061.8	99.7	1314	2	US-08-811-949-50
17	1061.8	99.7	1974	2	US-08-811-949-38
18	1057	99.2	1068	2	US-08-811-949-52
19	1057	99.2	1314	2	US-08-811-949-54
20	1056.4	99.2	1068	1	US-08-137-116-2
21	1056.2	99.2	1137	4	US-09-553-498-9
22	1056.2	99.2	1137	4	US-09-618-869-9
23	1055.4	99.1	1170	2	US-08-811-949-66
24	1054.8	99.0	1065	1	US-08-427-640-1
25	1053.8	98.9	1314	2	US-08-811-949-56
26	1050	98.6	1068	6	5223256-3
27	1046.8	98.3	1065	1	US-08-427-640-5

28	1041	97.7	1068	2	US-08-811-949-58
29	1035.4	97.2	1068	1	US-08-427-640-3
30	1033	97.0	1065	2	US-08-811-949-60
31	953.4	89.5	2497	6	5185259-2
32	924	86.8	1068	1	US-08-427-640-7
33	801	75.2	1163	2	US-08-558-269-5
34	801	75.2	1163	3	US-09-410-882-5
35	626.6	58.8	1727	6	5244676-4
36	625	58.7	1605	6	5244676-1
37	472	44.3	472	2	US-08-811-949-40
38	417	39.2	453	6	5200340-3
39	264.4	24.8	1724	6	5200340-5
40	170.8	16.0	1236	1	US-07-957-039A-7
41	170.8	16.0	1475	4	US-09-643-597-122
42	170.8	16.0	1475	4	US-09-480-884A-122
43	170.8	16.0	1475	4	US-09-542-615A-122
44	170.8	16.0	1475	4	US-09-606-421B-122
45	170.8	16.0	2301	6	5188829-2

#### ALIGNMENTS

#### RESULT 1

US-08-811-949-44  
; Sequence 44, Application US/08811949  
; Patent No. 5840533  
; GENERAL INFORMATION:  
; APPLICANT: NIWA, MINEO  
; APPLICANT: SAITO, YOSHIMASA  
; APPLICANT: SASAKI, HITOSHI  
; APPLICANT: HAYASHI, MASAKO  
; APPLICANT: NOTANI, JOUJI  
; APPLICANT: KOBAYASHI, MASAKAZU  
; TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR  
; NUMBER OF SEQUENCES: 67  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
; ADDRESS: P.C.  
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400  
; CITY: ARLINGTON  
; STATE: VA  
; COUNTRY: USA  
; ZIP: 22202

COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/811,949  
; FILING DATE: 05-MAR-1997  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: OBLON, NORMAN F.  
; REGISTRATION NUMBER: 24,618  
; REFERENCE/DOCKET NUMBER: 18-966-0  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 703-413-3000  
; TELEFAX: 703-413-2220  
; INFORMATION FOR SEQ ID NO: 44:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1068 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: circular  
; MOLECULE TYPE: DNA (genomic)  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 1..1065

US-08-811-949-44  
Query Match 100.0%; Score 1065; DB 2; Length 1068;





QY 661 GTGTGCTTCCCGCGGAGACTGAGCTGCGGAGTGTGAGCTCTCCGGC 720  
Db 1356 GTGTGCTTCCCGCGGAGACTGAGCTGCGGAGTGTGAGCTCTCCGGC 1415  
QY 721 TACGCAAGCATGAGGCTTGTCTCTTCTATTTCGAGCGGCTGAAGAGGCTCATGTC 780  
Db 1416 TACGCAAGCATGAGGCTTGTCTCTTCTATTTCGAGCGGCTGAAGAGGCTCATGTC 1475  
QY 781 AGACTGTACCCATCCAGCGCTGACATCAACAATTTACTTAAAGAAAGTCAACGAC 840  
Db 1476 AGACTGTACCCATCCAGCGCTGACATCAACAATTTACTTAAAGAAAGTCAACGAC 1535  
QY 841 AACATGCTGTGTGCTGGAGACACTCGAGCGCGGCGCCAGGCAAACTTGCACGAGCC 900  
Db 1536 AACATGCTGTGTGCTGGAGACACTCGAGCGCGGCGCCAGGCAAACTTGCACGAGCC 1595  
QY 901 TGCCAGGCGGATTCGGAGGCGCCCTGTGTGTCTGAACGATGCGCGCATGTTGGTG 960  
Db 1596 TGCCAGGCGGATTCGGAGGCGCCCTGTGTGTCTGAACGATGCGCGCATGTTGGTG 1655  
QY 961 GGCATCATCAGTGGGCGCTGGGCTGTGGACAGAGGATGTCGGGTGTGTACAAAG 1020  
Db 1656 GGCATCATCAGTGGGCGCTGGGCTGTGGACAGAGGATGTCGGGTGTGTACAAAG 1715  
QY 1021 GTTACCAACTACCTAGACTGGATTCTGTGACAAACATGCGACCGTGA 1065  
Db 1716 GTTACCAACTACCTAGACTGGATTCTGTGACAAACATGCGACCGTGA 1760

## RESULT 4

US-08-811-949-42  
; Sequence 42, Application US/08811949  
; Patent No. 5840533  
; GENERAL INFORMATION:  
; APPLICANT: NIWA, MINEO  
; APPLICANT: SAITO, YOSHIMASA  
; APPLICANT: SASAKI, HITOSHI  
; APPLICANT: HAYASHI, MASAKO  
; APPLICANT: NOTANI, JOUJI  
; APPLICANT: KOBAYASHI, MASAKAZU  
; TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR  
; NUMBER OF SEQUENCES: 67  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,  
; ADDRESSEE: P.C.  
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400  
; CITY: ARLINGTON  
; STATE: VA  
; COUNTRY: USA  
; ZIP: 22202  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/811,949  
; FILING DATE: 05-MAR-1997  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: OBLON, NORMAN F.  
; REGISTRATION NUMBER: 24,618  
; REFERENCE/DOCKET NUMBER: 18-966-0  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 703-413-3000  
; TELEFAX: 703-413-2220  
; INFORMATION FOR SEQ ID NO: 42:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2101 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: circular  
; MOLECULE TYPE: DNA (genomic)

FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 25...1710  
US-08-811-949-42  
Query Match 100.0%; Score 1065; DB 2; Length 2101;  
Best Local Similarity 100.0%; Pred. No. 1.8e-291;  
Matches 1065; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 TCTGAGGAAACAGTGTCTCTACTTTGGGAATGGGTACGCTACCGTGGGACGACAGC 60  
Db 649 TCTGAGGAAACAGTGTCTCTACTTTGGGAATGGGTACGCTACCGTGGGACGACAGC 708  
QY 61 CTACCCGAGTGGGTGCTCTCTGCTCCCGTGGAAATTCATGATTCCTGATAGGCAAGTT 120  
Db 709 CTACCCGAGTGGGTGCTCTCTGCTCCCGTGGAAATTCATGATTCCTGATAGGCAAGTT 768  
QY 121 TACACAGCACAGAACCCAGTGCAGGCACTGGGCTGGGCAAAACAATAATTAATCTGCGG 180  
Db 769 TACACAGCACAGAACCCAGTGCAGGCACTGGGCTGGGCAAAACAATAATTAATCTGCGG 828  
QY 181 AATCTGATGGGATGCCAAGCCCTGTGTGCAAGTGTGAAGAACCGGAGGCTGACGTGG 240  
Db 829 AATCTGATGGGATGCCAAGCCCTGTGTGCAAGTGTGAAGAACCGGAGGCTGACGTGG 888  
QY 241 GAGTACTGTGATGTGCCCTCTCTGCTCCACCTGGGCTGAGACAGTACAGCCCTCAG 300  
Db 889 GAGTACTGTGATGTGCCCTCTCTGCTCCACCTGGGCTGAGACAGTACAGCCCTCAG 948  
QY 301 TTTGCGATCAAGAGGAGGCTCTTTGCGCGACATCGCCTCCCAACCCCTGGCAGGCTGCCATC 360  
Db 949 TTTGCGATCAAGAGGAGGCTCTTTGCGCGACATCGCCTCCCAACCCCTGGCAGGCTGCCATC 1008  
QY 361 TTTGCGCAAGCACAGGAGGTCGCGCGGAGAGCGGTTCCTGTGCGGCGGCATCTCATCAGC 420  
Db 1009 TTTGCGCAAGCACAGGAGGTCGCGCGGAGAGCGGTTCCTGTGCGGCGGCATCTCATCAGC 1068  
QY 421 TCCTGTGGAATTCCTCTGCGCGCCACTGCTTCAGGAGAGGTTTCCGCGCCCAACCTG 480  
Db 1069 TCCTGTGGAATTCCTCTGCGCGCCACTGCTTCAGGAGAGGTTTCCGCGCCCAACCTG 1128  
QY 481 ACGGTGATCTTGGGCAAGACATACCGGCTGTCTTCCCTGGCGAGGAGGAGAAATTTGAA 540  
Db 1129 ACGGTGATCTTGGGCAAGACATACCGGCTGTCTTCCCTGGCGAGGAGGAGAAATTTGAA 1188  
QY 541 GTCGAAAAATACATTTGCTCATAGGAATTCGATGATGACACTTACGACAAATGACATTCGG 600  
Db 1189 GTCGAAAAATACATTTGCTCATAGGAATTCGATGATGACACTTACGACAAATGACATTCGG 1248  
QY 601 CTGCTGAGCTGAAATCGGATTCGTCCTGCTGCGCGAGGAGGAGGAGGCTGGTCCGCACT 660  
Db 1249 CTGCTGAGCTGAAATCGGATTCGTCCTGCTGCGCGAGGAGGAGGAGGCTGGTCCGCACT 1308  
QY 661 GTGTGCTTCCCGCGGAGGCTGACGCTGCGGAGTGGAGCGGAGTGTGAGCTCTCCGGC 720  
Db 1309 GTGTGCTTCCCGCGGAGGCTGACGCTGCGGAGTGGAGCGGAGTGTGAGCTCTCCGGC 1368  
QY 721 TACGGCAAGCATGAGGCTTGTCTCTTCTATTTCGAGCGGCTGAAGAGGCTCATGTC 780  
Db 1369 TACGGCAAGCATGAGGCTTGTCTCTTCTATTTCGAGCGGCTGAAGAGGCTCATGTC 1428  
QY 781 AGACTGTACCCATCCAGCGCTGACATCAACAATTTACTTAAAGAAAGTCAACGAC 840  
Db 1429 AGACTGTACCCATCCAGCGCTGACATCAACAATTTACTTAAAGAAAGTCAACGAC 1488  
QY 841 AACATGCTGTGTGCTGGAGACACTCGGAGCGGCGGCGCCAGGCAAACTTGCACGAGCC 900  
Db 1489 AACATGCTGTGTGCTGGAGACACTCGGAGCGGCGGCGCCAGGCAAACTTGCACGAGCC 1548  
QY 901 TGCCAGGCGGATTCGGAGGCGCCCTGTGTGTGTAACGATGCGCGCATGTTGGTG 960  
Db 1549 TGCCAGGCGGATTCGGAGGCGCCCTGTGTGTGTAACGATGCGCGCATGTTGGTG 1608

QY 961 GGCATCATCAGCTGGGGCTGGCTGTGGACAGAGAGATGTCCCGGGTGTGTACACAAAG 1020  
DB 1609 GGCATCATCAGCTGGGGCTGGCTGTGGACAGAGAGATGTCCCGGGTGTGTACACAAAG 1668  
QY 1021 GTTACCAACTACTAGACTGGATTCTGTGACACACATGCGACCGTGA 1065  
DB 1669 GTTACCAACTACTAGACTGGATTCTGTGACACACATGCGACCGTGA 1713

RESULT 5

US-08-119-512-3  
; Sequence 3, Application US/08119512  
; Patent No. 5498531  
; GENERAL INFORMATION:  
; APPLICANT: Jarrell, Kevin A.  
; TITLE OF INVENTION: INTRON MEDIATED RECOMBINANT TECHNIQUES  
; NUMBER OF INVENTIONS: AND REAGENTS  
; NUMBER OF SEQUENCES: 3  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: LAHIVE & COCKFIELD  
; STREET: 60 State Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: ASCII (text)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/119.512  
; FILING DATE: 10-SEP-1993  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Vincent, Matthew P.  
; REGISTRATION NUMBER: 36,709  
; REFERENCE/DOCKET NUMBER: HUI-008  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 227-7400  
; TELEFAX: (617) 227-5941  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2162 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: both  
; MOLECULE TYPE: other nucleic acid  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: 82..334  
; OTHER INFORMATION: /product= "Signal Sequence and  
; OTHER INFORMATION: Finger-like domain"  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: 335..447  
; OTHER INFORMATION: /product= "EGF-like domain"  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: 448..714  
; OTHER INFORMATION: /product= "Kringle-1 domain"  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: 715..972  
; OTHER INFORMATION: /product= "Kringle-2 domain"  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: 973..2162  
; OTHER INFORMATION: /product= "Catalytic domain"  
US-08-119-512-3

Query Match 100.0%; Score 1065; DB 1; Length 2162;  
Best Local Similarity 100.0%; Pred. No. 1.8e-291;

Matches 1065; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 TCTCAGGAAACAGTACTGTCTACTTTGGGAAATGGGTACGCTACCGTGGCAGCAGCAGC 60  
DB 706 TCTCAGGAAACAGTACTGTCTACTTTGGGAAATGGGTACGCTACCGTGGCAGCAGCAGC 765  
QY 61 CTACCCGAGTGGGTGCTCTCTGCTCCCGTGGAAATTCATGATCTCTGTATAGGCAAGTT 120  
DB 766 CTACCCGAGTGGGTGCTCTCTGCTCCCGTGGAAATTCATGATCTCTGTATAGGCAAGTT 825  
QY 121 TACACAGACAGAACCCAGTGGCCAGGCACTGGGCTGGGCAACATATATTTACTGCCGG 180  
DB 826 TACACAGACAGAACCCAGTGGCCAGGCACTGGGCTGGGCAACATATATTTACTGCCGG 885  
QY 181 AATCTGTATGGGATGCCAAGCCCTGTGTGCGACATCGCTGGAAGAACCCAGGCTCACGTGG 240  
DB 886 AATCTGTATGGGATGCCAAGCCCTGTGTGCGACATCGCTGGAAGAACCCAGGCTCACGTGG 945  
QY 241 GAGTACTGTATGTCCTCTCTGCTCCACCTGCGGCTGTGAGACAGTACAGCAGCCTCAG 300  
DB 946 GAGTACTGTATGTCCTCTCTGCTCCACCTGCGGCTGTGAGACAGTACAGCAGCCTCAG 1005  
QY 301 TTTCCGATCAAGAGGGCTCTTCCGCGACATCGCTCCACCCCTGGCAGGCTGCCATC 360  
DB 1006 TTTCCGATCAAGAGGGCTCTTCCGCGACATCGCTCCACCCCTGGCAGGCTGCCATC 1065  
QY 361 TTTGCCAAGCACAGGAGTGGCCGAGAGCGGTTCCTGTGCGGGGCGATCTCATCAGC 420  
DB 1066 TTTGCCAAGCACAGGAGTGGCCGAGAGCGGTTCCTGTGCGGGGCGATCTCATCAGC 1125  
QY 421 TCCTGTGAGATTCTCTGCGGCGCACTGTCTTCAGGAGAGGTTTCGCGCCCAACCACTG 480  
DB 1126 TCCTGTGAGATTCTCTGCGGCGCACTGTCTTCAGGAGAGGTTTCGCGCCCAACCACTG 1185  
QY 481 ACGGTGATCTTGGCGAGAAATACACCGGGTGGTCCCTGGCGAGGAGGAGAAATTTGAA 540  
DB 1186 ACGGTGATCTTGGCGAGAAATACACCGGGTGGTCCCTGGCGAGGAGGAGAAATTTGAA 1245  
QY 541 GTCGAAAAATACATTTGTCCATAGGAATTCGATGATGACACTTTACAGCAATGACATTTGCG 600  
DB 1246 GTCGAAAAATACATTTGTCCATAGGAATTCGATGATGACACTTTACAGCAATGACATTTGCG 1305  
QY 601 CTGCTGAGCTGAAATCGGAATTCGCTCCGCTGTGCCAGAGAGCAGCGTGGTCCGCACT 660  
DB 1306 CTGCTGAGCTGAAATCGGAATTCGCTCCGCTGTGCCAGAGAGCAGCGTGGTCCGCACT 1365  
QY 661 GTGTGCTTCCCGCGGAGCTGCGAGTCCGCGAGTGGACGAGTGTGAGCTTCCGGC 720  
DB 1366 GTGTGCTTCCCGCGGAGCTGCGAGTCCGCGAGTGGACGAGTGTGAGCTTCCGGC 1425  
QY 721 TACGGCAAGCATGAGGCTTTGTCTCTTCTATTTCGGAGCGGCTGAAGGAGGCTCATGTC 780  
DB 1426 TACGGCAAGCATGAGGCTTTGTCTCTTCTATTTCGGAGCGGCTGAAGGAGGCTCATGTC 1485  
QY 781 AGACTGTACCCATCCAGCCGCTGCACATCAACATTTACTTAACAGAACAGTCACCGAC 840  
DB 1486 AGACTGTACCCATCCAGCCGCTGCACATCAACATTTACTTAACAGAACAGTCACCGAC 1545  
QY 841 AACATGCTGTGCTGGAGACACTCGGAGCGGGGGCCCCAGGCAAACTTTCACAGCGCC 900  
DB 1546 AACATGCTGTGCTGGAGACACTCGGAGCGGGGGCCCCAGGCAAACTTTCACAGCGCC 1605  
QY 901 TGCAGGCGGATTCGGGAGCGCCCCCTGTGTGTGAACGATGGCCGATGACTTTGGTG 960  
DB 1606 TGCAGGCGGATTCGGGAGCGCCCCCTGTGTGTGAACGATGGCCGATGACTTTGGTG 1665  
QY 961 GGCATCATCAGCTGGGCGCTGGGCTGTGGACAGAGGATGTCCCGGGTGTGTACAAAG 1020  
DB 1666 GGCATCATCAGCTGGGCGCTGGGCTGTGGACAGAGGATGTCCCGGGTGTGTACAAAG 1725  
QY 1021 GTTACCAACTACTAGACTGGATTCTGTGACACACATGCGACCGTGA 1065  
DB 1726 GTTACCAACTACTAGACTGGATTCTGTGACACACATGCGACCGTGA 1770

RESULT 6  
US-08-488-015B-3  
; Sequence 3, Application US/08488015B  
; Patent No. 5780272  
; GENERAL INFORMATION:  
; APPLICANT: Jarrell, Kevin A.  
; TITLE OF INVENTION: INTRON-MEDIATED RECOMBINANT TECHNIQUES  
; TITLE OF INVENTION: AND REAGENTS  
; NUMBER OF SEQUENCES: 29  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Foley, Hoag & Eliot  
; STREET: One Post Office Square  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: ASCII (text)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/488.015B  
; FILING DATE: 07-JUN-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Vincent, Matthew P.  
; REGISTRATION NUMBER: 36,709  
; REFERENCE/DOCKET NUMBER: HUV-008.02  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 832-1000  
; TELEFAX: (617) 832-7000  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2162 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: both  
; MOLECULE TYPE: other nucleic acid  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: 82..334  
; OTHER INFORMATION: /product= "Signal Sequence and  
; Patent No. 5780272  
; OTHER INFORMATION: Finger-like domain"  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: 335..447  
; OTHER INFORMATION: /product= "EGF-like domain"  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: 448..714  
; OTHER INFORMATION: /product= "Kringle-1 domain"  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: 715..972  
; OTHER INFORMATION: /product= "Kringle-2 domain"  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: 973..2162  
; OTHER INFORMATION: /product= "Catalytic domain"  
US-08-488-015B-3

Query Match 100.0%; Score 1065; DB 1; Length 2162;  
Best Local Similarity 100.0%; Pred. No. 1.8e-291;  
Matches 1065; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db	Qy	766	121	826	181	886	241	946	301	1006	361	1066	421	1126	481	1186	541	1246	601	1306	661	1366	721	1426	781	1486	841	1546	901	1606	961	1666	1021	1726		
Db	Qy	CTCACCAGTGGGTCCTCTCCCTCCGCTGAAATTCATGATCCTGATAGCAAGTT	TACACAGCACAGAACCCAGTCCAGGCACTGGGCTGGGCAACATANTTACTCGGG	TACACAGCACAGAACCCAGTCCAGGCACTGGGCTGGGCAACATANTTACTCGGG	AATCCTGATGGGATGCCAAGCCTCTGTCGACGCTGCTGAAGAACCGCAGCTGAC	AATCCTGATGGGATGCCAAGCCTCTGTCGACGCTGCTGAAGAACCGCAGCTGAC	GAGTACTGTGATGTGCCCTCTCTCCACCTCGCGGCTGAGACAGTACAGCAGCCTC	GAGTACTGTGATGTGCCCTCTCTCCACCTCGCGGCTGAGACAGTACAGCAGCCTC	TTTCGCATCAAGAGGGCTCTTCGCGGACATGCCCTCCACCCCTGGCAGGCTGCCATC	TTTCGCATCAAGAGGGCTCTTCGCGGACATGCCCTCCACCCCTGGCAGGCTGCCATC	TTTCGCATCAAGAGGGCTCTTCGCGGACATGCCCTCCACCCCTGGCAGGCTGCCATC	TTTCGCATCAAGAGGGCTCTTCGCGGACATGCCCTCCACCCCTGGCAGGCTGCCATC	TTTCGCATCAAGAGGGCTCTTCGCGGACATGCCCTCCACCCCTGGCAGGCTGCCATC	TTTCGCATCAAGAGGGCTCTTCGCGGACATGCCCTCCACCCCTGGCAGGCTGCCATC	TTTCGCATCAAGAGGGCTCTTCGCGGACATGCCCTCCACCCCTGGCAGGCTGCCATC	TTTCGCATCAAGAGGGCTCTTCGCGGACATGCCCTCCACCCCTGGCAGGCTGCCATC	TTTCGCATCAAGAGGGCTCTTCGCGGACATGCCCTCCACCCCTGGCAGGCTGCCATC	TTTCGCATCAAGAGGGCTCTTCGCGGACATGCCCTCCACCCCTGGCAGGCTGCCATC	TTTCGCATCAAGAGGGCTCTTCGCGGACATGCCCTCCACCCCTGGCAGGCTGCCATC	TTTCGCATCAAGAGGGCTCTTCGCGGACATGCCCTCCACCCCTGGCAGGCTGCCATC	TTTCGCATCAAGAGGGCTCTTCGCGGACATGCCCTCCACCCCTGGCAGGCTGCCATC	TTTCGCATCAAGAGGGCTCTTCGCGGACATGCCCTCCACCCCTGGCAGGCTGCCATC	TTTCGCATCAAGAGGGCTCTTCGCGGACATGCCCTCCACCCCTGGCAGGCTGCCATC	TTTCGCATCAAGAGGGCTCTTCGCGGACATGCCCTCCACCCCTGGCAGGCTGCCATC	TTTCGCATCAAGAGGGCTCTTCGCGGACATGCCCTCCACCCCTGGCAGGCTGCCATC	TTTCGCATCAAGAGGGCTCTTCGCGGACATGCCCTCCACCCCTGGCAGGCTGCCATC	TTTCGCATCAAGAGGGCTCTTCGCGGACATGCCCTCCACCCCTGGCAGGCTGCCATC	TTTCGCATCAAGAGGGCTCTTCGCGGACATGCCCTCCACCCCTGGCAGGCTGCCATC	TTTCGCATCAAGAGGGCTCTTCGCGGACATGCCCTCCACCCCTGGCAGGCTGCCATC	TTTCGCATCAAGAGGGCTCTTCGCGGACATGCCCTCCACCCCTGGCAGGCTGCCATC	TTTCGCATCAAGAGGGCTCTTCGCGGACATGCCCTCCACCCCTGGCAGGCTGCCATC	TTTCGCATCAAGAGGGCTCTTCGCGGACATGCCCTCCACCCCTGGCAGGCTGCCATC	TTTCGCATCAAGAGGGCTCTTCGCGGACATGCCCTCCACCCCTGGCAGGCTGCCATC	TTTCGCATCAAGAGGGCTCTTCGCGGACATGCCCTCCACCCCTGGCAGGCTGCCATC	TTTCGCATCAAGAGGGCTCTTCGCGGACATGCCCTCCACCCCTGGCAGGCTGCCATC	TTTCGCATCAAGAGGGCTCTTCGCGGACATGCCCTCCACCCCTGGCAGGCTGCCATC

RESULT 7  
US-08-488-015B-25  
; Sequence 25, Application US/08488015B  
; Patent No. 5780272  
; GENERAL INFORMATION:

Qy 1 TCTGAGGAAACAGTACTTCTGGGATGGTCCAGCTACCGTGGCAGCCAGC 60  
Db 706 TCTGAGGAAACAGTACTTCTGGGATGGTCCAGCTACCGTGGCAGCCAGC 765  
Qy 61 CTCACCGAGTGGGTCCTCTCCCTCCGCTGAAATTCATGATCCTGATAGCAAGTT 120



APPLICANT: Jarrell, Kevin A.  
TITLE OF INVENTION: INTRON-MEDIATED RECOMBINANT TECHNIQUES  
NUMBER OF SEQUENCES: 29  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Foley, Hoag & Eliot  
STREET: One Post Office Square  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII (text)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/488,015B  
FILING DATE: 07-JUN-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Vincent, Matthew P.  
REGISTRATION NUMBER: 36,709  
REFERENCE/DOCKET NUMBER: HUV-008.02  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 832-1000  
TELEFAX: (617) 832-7000  
INFORMATION FOR SEQ ID NO: 25:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2162 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
US-08-488-015B-25

Query Match 100.0%; Score 1065; DB 1; Length 2162;  
Best Local Similarity 100.0%; Pred. No. 1.8e-291;  
Matches 1065; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCTGAGGAAACAGTACTGCTTCTGGGAATGGGTGAGTCCGCTGAGCGGACGACAGC 60  
DB |  
706 TCTGAGGAAACAGTACTGCTTCTGGGAATGGGTGAGTCCGCTGAGCGGACGACAGC 765  
QY 61 CTCACCGAGTGGGTGCTCTGCTCCCTGCTGGAATTCATGATCTCTGATAGGCAAGTT 120  
DB |  
766 CTCACCGAGTGGGTGCTCTGCTCCCTGCTGGAATTCATGATCTCTGATAGGCAAGTT 825  
QY 121 TACACAGACAGAACCCAGTCCGAGGACTGGGCTGGGCAACATAATTAATCTGCCGG 180  
DB |  
826 TACACAGACAGAACCCAGTCCGAGGACTGGGCTGGGCAACATAATTAATCTGCCGG 885  
QY 181 AATCTGATGGGATGCCAAGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240  
DB |  
886 AATCTGATGGGATGCCAAGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 945  
QY 241 GAGTACTGTGATGTCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300  
DB |  
946 GAGTACTGTGATGTCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1005  
QY 301 TTTGTCATCAAGAGGGCTCTTGGCCAGATCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCC 360  
DB |  
1006 TTTGTCATCAAGAGGGCTCTTGGCCAGATCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCC 1065  
QY 361 TTTGTCATCAAGAGGGCTCTTGGCCAGATCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCC 420  
DB |  
1066 TTTGTCATCAAGAGGGCTCTTGGCCAGATCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCC 1125  
QY 421 TCCTGCTGGATTCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480  
DB |  
1126 TCCTGCTGGATTCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1185  
QY 481 ACGGTGATCTTGGGAGAACATACCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540  
DB |

DB 1186 ACGGTGATCTTGGGAGAACATACCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1245  
QY 541 GTCGAAAATACATTTGTCATAGGAATTCGATGATGACACTTACGACAAATGACATTCGG 600  
DB |  
1246 GTCGAAAATACATTTGTCATAGGAATTCGATGATGACACTTACGACAAATGACATTCGG 1305  
QY 601 CTGCTGAGCTGAAATCGGATTCCTCCGCTGCTCCGAGGAGAGCAGCGTGGTCCGCACT 660  
DB |  
1306 CTGCTGAGCTGAAATCGGATTCCTCCGCTGCTCCGAGGAGAGCAGCGTGGTCCGCACT 1365  
QY 661 GTGTGCTTCCCGGGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720  
DB |  
1366 GTGTGCTTCCCGGGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1425  
QY 721 TACGCAAGCATGAGGCTTGTCTCTCTTCTTATTCGAGCGGCTGAAAGGAGGCTCATGTC 780  
DB |  
1426 TACGCAAGCATGAGGCTTGTCTCTCTTCTTATTCGAGCGGCTGAAAGGAGGCTCATGTC 1485  
QY 781 AGACTGATCCCATCCAGCGGCTGCATCAACAATTTACTTAAACAGAAAGTCAACGAC 840  
DB |  
1486 AGACTGATCCCATCCAGCGGCTGCATCAACAATTTACTTAAACAGAAAGTCAACGAC 1545  
QY 841 AACATGCTGTGCTGAGGACACTCGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 900  
DB |  
1546 AACATGCTGTGCTGAGGACACTCGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 1605  
QY 901 TGCCAGGCGGATTCGAGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 960  
DB |  
1606 TGCCAGGCGGATTCGAGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1665  
QY 961 GGCATCATCAGTGGGCGCTGGGCTGTGGACAGAAAGATGTCGCGGTGTGTACACAAAG 1020  
DB |  
1666 GGCATCATCAGTGGGCGCTGGGCTGTGGACAGAAAGATGTCGCGGTGTGTACACAAAG 1725  
QY 1021 GTTACCAACTACTAGTGGATTCGTGACAACTGCGACGCGGTGA 1065  
DB |  
1726 GTTACCAACTACTAGTGGATTCGTGACAACTGCGACGCGGTGA 1770

## RESULT 8

US-08-811-949-64  
; Sequence 64, Application US/08811949  
; Patent No. 5840533  
; GENERAL INFORMATION:  
; APPLICANT: NIWA, MINEO  
; APPLICANT: SAITO, YOSHIMASA  
; APPLICANT: SASAKI, HITOSHI  
; APPLICANT: HAYASHI, MASAKO  
; APPLICANT: NOTANI, JOUJI  
; APPLICANT: KOBAYASHI, MASAKAZU  
; TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR  
; NUMBER OF SEQUENCES: 67  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
; ADDRESSEE: P. C.  
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400  
; CITY: ARLINGTON  
; STATE: VA  
; COUNTRY: USA  
; ZIP: 22202  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/811,949  
; FILING DATE: 05-MAR-1997  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: OBLON, NORMAN F.  
; REGISTRATION NUMBER: 24,618  
; REFERENCE/DOCKET NUMBER: 18-966-0

## TELECOMMUNICATION INFORMATION:

TELEPHONE: 703-413-3000  
TELEFAX: 703-413-2220  
INFORMATION FOR SEQ ID NO: 64:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1170 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: circular  
MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..1167  
US-08-811-949-64

Query Match 99.8%; Score 1063.4; DB 2; Length 1170;

Best Local Similarity 99.9%; Pred. No. 4e-291;  
Matches 1064; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TCTGAGGAAACAGTACTGCTTCTTGGGAATGGCTCAGCTTACCGTGGCAGCAGCAGC 60  
DB 106 TCTGAGGAAACAGTACTGCTTCTTGGGAATGGCTCAGCTTACCGTGGCAGCAGCAGC 165  
QY 61 CTCACCGAGTGGGTGCTCTCTGCTCCCTCCGTTGGAAATCCATGATCCTGATAGGCAAGTT 120  
DB 166 CTCACCGAGTGGGTGCTCTCTGCTCCCTCCGTTGGAAATCCATGATCCTGATAGGCAAGTT 225  
QY 121 TACACAGACAGAACCCAGTCCGAGCAGTGGGCTGGGCAACATAATTAATCTGCGG 180  
DB 226 TACACAGACAGAACCCAGTCCGAGCAGTGGGCTGGGCAACATAATTAATCTGCGG 285  
QY 181 AATCTGATGGGATGCCAAGCCCTGGTGGCAGCTGCTGAAGAACCGCAGCTGACGTGG 240  
DB 286 AATCTGATGGGATGCCAAGCCCTGGTGGCAGCTGCTGAAGAACCGCAGCTGACGTGG 345  
QY 241 GAGTACTGTGATGCTCTCTCTGCTCCACCTGCGGCTTGAGACAGTACAGCCAGCTCAG 300  
DB 346 GAGTACTGTGATGCTCTCTCTGCTCCACCTGCGGCTTGAGACAGTACAGCCAGCTCAG 405  
QY 301 TTTCGATCAAGAGGGGCTCTTCCGACAGTCCGCTCCACCTGCGAGCTGCGCATC 360  
DB 406 TTTCGATCAAGAGGGGCTCTTCCGACAGTCCGCTCCACCTGCGAGCTGCGCATC 465  
QY 361 TTTCGACAGCAGAGGTGCGCGGAGAGGGTTCCTGTCGGGGGATCTATCATCAGC 420  
DB 466 TTTCGACAGCAGAGGTGCGCGGAGAGGGTTCCTGTCGGGGGATCTATCATCAGC 525  
QY 421 TCCTGCTGGATCTCTCTGCGGCCACTGCTTCCAGGAGAGGTTTCCGCCCCACCATCTG 480  
DB 526 TCCTGCTGGATCTCTCTGCGGCCACTGCTTCCAGGAGAGGTTTCCGCCCCACCATCTG 585  
QY 481 ACGGTGATCTGGGAGAACATACCGGTGCTCCCTGCGAGGAGGAGCAGAAATTTGAA 540  
DB 586 ACGGTGATCTGGGAGAACATACCGGTGCTCCCTGCGAGGAGGAGCAGAAATTTGAA 645  
QY 541 GTCCGAAATATATTGTCATAAGAAATTCGATGATGACATTACGACAAATGACATTCGC 600  
DB 646 GTCCGAAATATATTGTCATAAGAAATTCGATGATGACATTACGACAAATGACATTCGC 705  
QY 601 CTGCTGACGTGAATTCGGATTCGCTCCGCTGTGCCCAGGAGAGCAGCGTGGTCCGCACT 660  
DB 706 CTGCTGACGTGAATTCGGATTCGCTCCGCTGTGCCCAGGAGAGCAGCGTGGTCCGCACT 765  
QY 661 GTGTGCTCTCCCGCGGAGCTGAGCTGCGGAGTGGAGCGGAGTGGAGCTTCCGGC 720  
DB 766 GTGTGCTCTCCCGCGGAGCTGAGCTGCGGAGTGGAGCGGAGTGGAGCTTCCGGC 825  
QY 721 TACGCAAGCATGAGCGCTTGTCTCTTCTTATTCGGAGCGGCTGAAGAGGCTCATGTC 780  
DB 826 TACGCAAGCATGAGCGCTTGTCTCTTCTTATTCGGAGCGGCTGAAGAGGCTCATGTC 885  
QY 781 AGACTGTACCCATCCAGCGCTGCAATCAACAATTTACTTTAACAAGACAGTCAACGAC 840

DB 886 AGACTGTACCCATCCAGCGCTGCACATCAACAATTTACTTAACAAGACAGTCAACGAC 945  
QY 841 AACATGCTGTGCTGGAGACATCCGAGCGGGGGCCCCAGGAAACTTGCACGAGCC 900  
DB 946 AACATGCTGTGCTGGAGACATCCGAGCGGGGGCCCCAGGAAACTTGCACGAGCC 1005  
QY 901 TGCCAGGGCGATTCCGGAGGCCCCCTGCTGTCTGAACGATGCGCCGATGCTTTGGTG 960  
DB 1006 TGCCAGGGCGATTCCGGAGGCCCCCTGCTGTCTGAACGATGCGCCGATGCTTTGGTG 1065  
QY 961 GGCATCATCAGCTGGGCGCTGGGCTGTGGACAGAGGATGTCGCGGTGTGTACAAAG 1020  
DB 1066 GGCATCATCAGCTGGGCGCTGGGCTGTGGACAGAGGATGTCGCGGTGTGTACAAAG 1125  
QY 1021 GTTACCAACTACTAGACTGGATTCTGTCGACAAATGCGACCGTGA 1065  
DB 1126 GTTACCAACTACTAGACTGGATTCTGTCGACAAATGCGACCGTGA 1170

## RESULT 9

US-08-811-949-48  
Sequence 48, Application US/08811949  
Patent No. 5840533  
GENERAL INFORMATION:  
APPLICANT: NIWA, MINEO  
APPLICANT: SAITO, YOSHIMASA  
APPLICANT: SAKI, HITOSHI  
APPLICANT: HAYASHI, MASAKO  
APPLICANT: NOTANI, JOUJI  
APPLICANT: KOBAYASHI, MASAKAZU  
TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR  
NUMBER OF SEQUENCES: 67  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,  
ADDRESS: P. C.  
STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400  
CITY: ARLINGTON  
STATE: VA  
COUNTRY: USA  
ZIP: 22202  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/811,949  
FILING DATE: 05-MAR-1997  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: OBLON, NORMAN F.  
REGISTRATION NUMBER: 24,618  
REFERENCE/DOCKET NUMBER: 18-966-0  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-413-3000  
TELEFAX: 703-413-2220  
INFORMATION FOR SEQ ID NO: 48:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1314 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: circular  
MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..1311  
US-08-811-949-48

Query Match 99.8%; Score 1063.4; DB 2; Length 1314;

Best Local Similarity 99.9%; Pred. No. 4.2e-291;  
Matches 1064; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TCTGAGGAAACAGTACTGCTTCTTGGGAATGGGTGAGCTTACCGTGGCAGCAGCAGC 60

Db 250 TCTGAGGAAACAGTGACTGCTACTTTGGGAATGGGTACGCTACCGTGGCAGCGCACAGC 309  
Qy 61 CTCACCGAGTCGGGTCCCTCTCTCCCTCCGTCGGAATTCATGATCCTGATAGCAAGGTT 120  
Db 310 CTCACCGAGTCGGGTCCCTCTCTCCCTCCGTCGGAATTCATGATCCTGATAGCAAGGTT 369  
Qy 121 TACACAGCAGAACCCAGTCCAGGACCTGGGCTGGGCAACATAATTAATCTAGTCCGG 180  
Db 370 TACACAGCAGAACCCAGTCCAGGACCTGGGCTGGGCAACATAATTAATCTAGTCCGG 429  
Qy 181 AATCCTGATGGGATGCCAAGCCCTGCTGTCACAGTCTGAAGAACCGCAGGCTGACGTGG 240  
Db 430 AATCCTGATGGGATGCCAAGCCCTGCTGTCACAGTCTGAAGAACCGCAGGCTGACGTGG 489  
Qy 241 GAGTACTGTGATGTCCCTCTCTCTCCCTCCAGTCCGCGCTGAGACAGTACAGCCAGCCTCAG 300  
Db 490 GAGTACTGTGATGTCCCTCTCTCTCCCTCCAGTCCGCGCTGAGACAGTACAGCCAGCCTCAG 549  
Qy 301 TTTCGCATCAAGGAGGCTCTTCGCGGACATCGCTCCAGCCCTGGGAGGCTGCCATC 360  
Db 550 TTTCGCATCAAGGAGGCTCTTCGCGGACATCGCTCCAGCCCTGGGAGGCTGCCATC 609  
Qy 361 TTTCGCAAGCACAGGAGTCGCGCGGAGAGCGGTTCTGTGCGGGGATACTCATCAGC 420  
Db 610 TTTCGCAAGCACAGGAGTCGCGCGGAGAGCGGTTCTGTGCGGGGATACTCATCAGC 669  
Qy 421 TCCTGCTGGAATTCCTCTGCGCGCCCACTGCTTCCAGGAGAGGTTTCCGCGCCCAACACCTG 480  
Db 670 TCCTGCTGGAATTCCTCTGCGCGCCCACTGCTTCCAGGAGAGGTTTCCGCGCCCAACACCTG 729  
Qy 481 ACGGTGATCTTGGGAGAACATACACGGGTGTCCTGCGGAGGAGGAGAGCAAAATTTGAA 540  
Db 730 ACGGTGATCTTGGGAGAACATACACGGGTGTCCTGCGGAGGAGGAGGAGCAAAATTTGAA 789  
Qy 541 GTCCGAAATACATTTGCTCAATGAATTCGATGATGACACTTACGACAAATGACATTCGG 600  
Db 790 GTCCGAAATACATTTGCTCAATGAATTCGATGATGACACTTACGACAAATGACATTCGG 849  
Qy 601 CTGCTGACGTGAATTCGGATTCGTCGCGCTGTGCGCGAGGAGCAGCGGTGTCGCACT 660  
Db 850 CTGCTGACGTGAATTCGGATTCGTCGCGCTGTGCGCGAGGAGCAGCGGTGTCGCACT 909  
Qy 661 GTGTGCTTCCCGCGGAGACTGAGCTGCGGAGCTGGAAGAGTGTGAGCTCTCCGCG 720  
Db 910 GTGTGCTTCCCGCGGAGACTGAGCTGCGGAGCTGGAAGAGTGTGAGCTCTCCGCG 969  
Qy 721 TACGCGAGCATGAGGCTTGTCTCTCTTCTATTTCGAGCGGCTGAAGAGGCTCATGTC 780  
Db 970 TACGCGAGCATGAGGCTTGTCTCTCTTCTATTTCGAGCGGCTGAAGAGGCTCATGTC 1029  
Qy 781 AGACTGTACCCATCCAGCGCTGCATCAACAATTTACTTTACAGAAACAGTCAACGAC 840  
Db 1030 AGACTGTACCCATCCAGCGCTGCATCAACAATTTACTTTACAGAAACAGTCAACGAC 1089  
Qy 841 AACATGCTGTGTGAGAGACTCGGAGCGCGGCGCCAGGCAAACTTGACAGCGCC 900  
Db 1090 AACATGCTGTGTGAGAGACTCGGAGCGCGGCGCCAGGCAAACTTGACAGCGCC 1149  
Qy 901 TGCCAGGCGGATTCGGGAGGCGCCCTGTGTGTCTGAACGATGCGCGCATGACTTTGGTG 960  
Db 1150 TGCCAGGCGGATTCGGGAGGCGCCCTGTGTGTCTGAACGATGCGCGCATGACTTTGGTG 1209  
Qy 961 GGCAATCATCAGTGGGCGCTGGGCTGTGACAGAGGATGTCGCGGTGTGTACACAAAG 1020  
Db 1210 GGCAATCATCAGTGGGCGCTGGGCTGTGACAGAGGATGTCGCGGTGTGTACACAAAG 1269  
Qy 1021 GTTACCAACTACCTAGATCGGATTCGTGACAACTGCGACCGTGA 1065  
Db 1270 GTTACCAACTACCTAGATCGGATTCGTGACAACTGCGACCGTGA 1314

RESULT 10

5200340-1  
; Patent No. 5200340  
; APPLICANT: FOSTER, DONALD C.; MULVIHILL, EILEEN R.; O'HARA,  
; PATRICK J.; PINGEL, KURT; YOSHITAKE, SHINJI  
; TITLE OF INVENTION: THROMBIN-ACTIVATED TISSUE PLASMINOGEN  
; ACTIVATORS  
; NUMBER OF SEQUENCES: 34  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/53,412  
; FILING DATE: 22-MAY-1987  
; SEQ ID NO:1:  
; LENGTH: 1738  
5200340-1

Query Match 99.8%; Score 1063.4; DB 6; Length 1738;  
Best Local Similarity 99.9%; Pred. No. 4.8e-291;

Matches 1064; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 TCTGAGGAAACAGTGACTGCTACTTTGGGAATGGGTACGCTACCGTGGCAGCGCACAGC 60  
Db 640 TCTGAGGAAACAGTGACTGCTACTTTGGGAATGGGTACGCTACCGTGGCAGCGCACAGC 699  
Qy 61 CTCACCGAGTCGGGTCCCTCTCTCCCTCCGTCGGAATTCATGATCCTGATAGCAAGGTT 120  
Db 700 CTCACCGAGTCGGGTCCCTCTCTCCCTCCGTCGGAATTCATGATCCTGATAGCAAGGTT 759  
Qy 121 TACACAGCAGAACCCAGTCCCGAGGCACTGGGCTTGGGCAACATAATTAATCTAGTCCGG 180  
Db 760 TACACAGCAGAACCCAGTCCCGAGGCACTGGGCTTGGGCAACATAATTAATCTAGTCCGG 819  
Qy 181 AATCCTGATGGGATGCCAAGCCCTGCTGTCGCGGAAATTCATGATCCTGATAGCAAGGTT 240  
Db 820 AATCCTGATGGGATGCCAAGCCCTGCTGTCGCGGAAATTCATGATCCTGATAGCAAGGTT 879  
Qy 241 GAGTACTGTGATGTCCCTCTCTCCCTCCAGTCCGCGCTGAGACAGTACAGCCAGCCTCAG 300  
Db 880 GAGTACTGTGATGTCCCTCTCTCCCTCCAGTCCGCGCTGAGACAGTACAGCCAGCCTCAG 939  
Qy 301 TTTCGCATCAAGGAGGCTCTTCGCGGACATCGCTCCAGCCCTGGCAGGCTGCCATC 360  
Db 940 TTTCGCATCAAGGAGGCTCTTCGCGGACATCGCTCCAGCCCTGGCAGGCTGCCATC 999  
Qy 361 TTTCGCAAGCACAGGAGTCGCGCGGAGCGGTTCTGTGCGGGGATACTCATCAGC 420  
Db 1000 TTTCGCAAGCACAGGAGTCGCGCGGAGCGGTTCTGTGCGGGGATACTCATCAGC 1059  
Qy 421 TCCTGCTGGAATTCCTCTGCGCGCCACTGCTTCCAGGAGGTTTCCGCGCCCAACACTG 480  
Db 1060 TCCTGCTGGAATTCCTCTGCGCGCCACTGCTTCCAGGAGGTTTCCGCGCCCAACACTG 1119  
Qy 481 ACGGTGATCTTGGGCAACAATACCGGTGCTCCCTGCGGAGGAGGAGCAAAATTTGAA 540  
Db 1120 ACGGTGATCTTGGGCAACAATACCGGTGCTCCCTGCGGAGGAGGAGCAAAATTTGAA 1179  
Qy 541 GTCGAAAAATACATTTGTCATAAGGAATTCGATGATGACACTTACGACAAATGACATTCG 600  
Db 1180 GTCGAAAAATACATTTGTCATAAGGAATTCGATGATGACACTTACGACAAATGACATTCG 1239  
Qy 601 CTGCTGACGTGAAATCGGAATTCCTCCGCTGTGTCGCGGAGGAGGAGGAGGAGGAGTTCG 660  
Db 1240 CTGCTGACGTGAAATCGGAATTCCTCCGCTGTGTCGCGGAGGAGGAGGAGGAGGAGTTCG 1299  
Qy 661 GTGTGCTTCCCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 720  
Db 1300 GTGTGCTTCCCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1359  
Qy 721 TACGCGAAGCATGAGGCTTGTCTCTCTTCTATTTCGAGGCGGCTGAAGAGGCTCATGTC 780  
Db 1360 TACGCGAAGCATGAGGCTTGTCTCTCTTCTATTTCGAGGCGGCTGAAGAGGCTCATGTC 1419  
Qy 781 AGACTGTACCCATCCAGCGCTGCATCAACAATTTACTTTACAGAAACAGTCAACGAC 840  
Db 1420 AGACTGTACCCATCCAGCGCTGCATCAACAATTTACTTTACAGAAACAGTCAACGAC 1479

QY 841 AACATGCTGTGTCTGGAGACACTCGGAGCGGCGGCCAGGCAAACTTGCACGAGCC 900  
DB 1480 AACATGCTGTGTCTGGAGACACTCGGAGCGGCGGCCAGGCAAACTTGCACGAGCC 1539  
QY 901 TGCCAGGCGGATTCGGGAGGCCCTTGTGTGTCTGAACGATGCCCGCATGACTTTGGTG 960  
DB 1540 TGCCAGGCGGATTCGGGAGGCCCTTGTGTGTCTGAACGATGCCCGCATGACTTTGGTG 1599  
QY 961 GGCATCATCACTGGGGCTGGGCTGTGGACAGAGGATGTCCGGGTGTGTACACAAAG 1020  
DB 1600 GGCATCATCACTGGGGCTGGGCTGTGGACAGAGGATGTCCGGGTGTGTACACAAAG 1659  
QY 1021 GTTACCAACTACCTAGACTGGATTGCTGACAAACATGCGACCGTGA 1065  
DB 1660 GTTACCAACTACCTAGACTGGATTGCTGACAAACATGCGACCGTGA 1704

## RESULT 11

US-08-883-795A-39  
; Sequence 39, Application US/08883795A  
; Patent No. 5985607  
; GENERAL INFORMATION:  
; APPLICANT: Delcuve, Genevieve  
; APPLICANT: Awang, Gregor  
; TITLE OF INVENTION: Recombinant DNA Molecules and Expression  
; TITLE OF INVENTION: Vectors for Tissue Plasminogen Activator  
; NUMBER OF SEQUENCES: 39  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: BERESKIN & PARR  
; STREET: 40 King Street West  
; CITY: Toronto  
; STATE: Ontario  
; COUNTRY: Canada  
; ZIP: M5H 3Y2  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/883.795A  
; FILING DATE: 27-JUN-1997  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Gravelle, Micheline  
; REGISTRATION NUMBER: 40,261  
; REFERENCE/DOCKET NUMBER: 7841-062  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (416) 364-7311  
; TELEFAX: (416) 361-1398  
; INFORMATION FOR SEQ ID NO: 39:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1955 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; ORIGINAL SOURCE:  
; ORGANISM: Homo sapiens

US-08-883-795A-39

Query Match 99.8%; Score 1063.4; DB 2; Length 1955;  
Best Local Similarity 99.9%; Pred. No. 5e-291;  
Matches 1064; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 TCTGAGGAAACAGTGAAGTCTTCTGGGAAATGGGTGAGCTACCGTGGCAACGACAGC 60  
DB 715 TCTGAGGAAACAGTGAAGTCTTCTGGGAAATGGGTGAGCTACCGTGGCAACGACAGC 774  
QY 61 CTCACCGAGTGGGTGCTCTGCTCCCTGGGAAATTCATGATCCTGATAGGCAAGTT 120  
DB 775 CTCACCGAGTGGGTGCTCTGCTCCCTGGGAAATTCATGATCCTGATAGGCAAGTT 834

QY 121 TACACAGCACAGAAACCCAGTCCAGGCACTGGGCTGGGCAAAACATAATTAATCTGCGG 180  
DB 835 TACACAGCACAGAAACCCAGTCCAGGCACTGGGCTGGGCAAAACATAATTAATCTGCGG 894  
QY 181 AATCTCTGATGGGATGCCAAGCCCTGGTGCACGTGCTGAAGAACCGCAGGCTGACGTGG 240  
DB 895 AATCTCTGATGGGATGCCAAGCCCTGGTGCACGTGCTGAAGAACCGCAGGCTGACGTGG 954  
QY 241 GAGTACTGTGATGTCCCTCTCTCTCCACTGCGGGCTGAGACAGTACAGCAGCCTCAG 300  
DB 955 GAGTACTGTGATGTCCCTCTCTCTCCACTGCGGGCTGAGACAGTACAGCAGCCTCAG 1014  
QY 301 TTTTCGCATCAAGAGAGGCTCTTTGCGCGACATCGCTCCACCCCTGGCAGGCTGCCATC 360  
DB 1015 TTTTCGCATCAAGAGAGGCTCTTTGCGCGACATCGCTCCACCCCTGGCAGGCTGCCATC 1074  
QY 361 TTTTCCCAAGCACAGGAGGTGCGCCCGAGAGCGGTTCCTGTGCGGGGGGCATATCATCAGC 420  
DB 1075 TTTTCCCAAGCACAGGAGGTGCGCCCGAGAGCGGTTCCTGTGCGGGGGGCATATCATCAGC 1134  
QY 421 TCCTGTGTGATTTCTCTGCGCGCCACTGCTTCAGAGAGGTTTTCGCGCCGCCACCTG 480  
DB 1135 TCCTGTGTGATTTCTCTGCGCGCCACTGCTTCAGAGAGGTTTTCGCGCCGCCACCTG 1194  
QY 481 ACGGTGATCTTTGGGCAACATACCGGGTGGTCCCTGGCGAGGAGGAGCAAGAAATTTGAA 540  
DB 1195 ACGGTGATCTTTGGGCAACATACCGGGTGGTCCCTGGCGAGGAGGAGCAAGAAATTTGAA 1254  
QY 541 GTCGAAAAATACATTTGTCATAAGGAATTCGATGATGACACTTACGCAATGACATTTGCG 600  
DB 1255 GTCGAAAAATACATTTGTCATAAGGAATTCGATGATGACACTTACGCAATGACATTTGCG 1314  
QY 601 CTGCTGCAGCTGAAATCGGATTCGTCCCGCTGTGCCAGAGAGCAGCGTGGTCCGCACT 660  
DB 1315 CTGCTGCAGCTGAAATCGGATTCGTCCCGCTGTGCCAGAGAGCAGCGTGGTCCGCACT 1374  
QY 661 GTGTGCTTCCCGCGGAGACCTGACGTCCCGGACTGGACGGAGTGTGAGCTCTCCGCG 720  
DB 1375 GTGTGCTTCCCGCGGAGACCTGACGTCCCGGACTGGACGGAGTGTGAGCTCTCCGCG 1434  
QY 721 TACGGCAAGCATGAGGCTTTGCTCTCTTTCTATTTCGAGCGGCTGAAGAGGCTCATGTC 780  
DB 1435 TACGGCAAGCATGAGGCTTTGCTCTCTTTCTATTTCGAGCGGCTGAAGAGGCTCATGTC 1494  
QY 781 AGACTGTACCCATCCAGCCGCTGCACATCAACATTTACTTAACAGAACAGTCCAGC 840  
DB 1495 AGACTGTACCCATCCAGCCGCTGCACATCAACATTTACTTAACAGAACAGTCCAGC 1554  
QY 841 AACATGCTGTGTCTGGAGACACTCGGAGCGGCGGCCAGGCAAACTTGCACGAGCC 900  
DB 1555 AACATGCTGTGTCTGGAGACACTCGGAGCGGCGGCCAGGCAAACTTGCACGAGCC 1614  
QY 901 TGCCAGGCGGATTCGGGAGGCCCTTGTGTGTGTAACGATGCCCGCATGACTTTGGTG 960  
DB 1615 TGCCAGGCGGATTCGGGAGGCCCTTGTGTGTGTAACGATGCCCGCATGACTTTGGTG 1674  
QY 961 GGCATCATCACTGGGGCTGGGCTGTGGACAGAGGATGTCCGGGTGTGTACACAAAG 1020  
DB 1675 GGCATCATCACTGGGGCTGGGCTGTGGACAGAGGATGTCCGGGTGTGTACACAAAG 1734  
QY 1021 GTTACCAACTACCTAGACTGGATTGCTGACAAACATGCGACCGTGA 1065  
DB 1735 GTTACCAACTACCTAGACTGGATTGCTGACAAACATGCGACCGTGA 1779

## RESULT 12

5344773-1  
; Patent No. 5344773  
; APPLICANT: WEI, CHA-MER; HSIUNG, NANCY; REDDY, VERMURI B.;  
; LEWONTT, JEFFREY F.; DACKOWSKI, WILLIAM; DOUGLAS, RICHARD;  
; COLE, EDWARD S.; PURCELL JR., RICHARD D.; LAU, DAVID TAI-YUI  
; TITLE OF INVENTION: HUMAN UTERINE TISSUE PLASMINOGEN

:ACTIVATOR PRODUCED BY RECOMBINANT DNA

: NUMBER OF SEQUENCES: 6

NUMBER OF REFERENCES: 0  
: CURRENT APPLICATION DATA:

CONCURRENT AFFIDAVIT DATA:  
APPLICATION NUMBER: IIS/07/782 686

REFLECTION NUMBER: 03  
EFFECTING DATE: 01-OCT-1988

FILED DATE: 01-01-1983  
PRIOR APPLICATION DATA.

PRIOR AFFILIATION DATA:  
APPLICATION NUMBER 656

APPLICANT  
FILING DATE

FILING DATE:  
; GEO TO NO. 1.

NOTES:

Query Match 99.8%; Score 1063.4; DB 6; Length 2457;

Query Match 33.8%; SCORE 1063.4; DB 6; Length 2437;  
Best Local Similarity 99.9%; Pred. No. 5.5e-291;

Best Local Similarity 99.9%; Pred. NO. 3.3E-291;  
Matches 1064: Conservative 0: Mismatches 1: Indels 0: Gaps 0:

QY	1	TCTGAGGAAACAGTGA	CTGCTGATCTTTGGGAATGGGTACGCTTACCGTGGCACGCCACGC	60
Db	638	TCTGAGGAAACAGTGA	CTGCTACTTTGGGAATGGGTACGCTTACCGTGGCACGCCACGC	697
QY	61	CTCACCGAGTCGGTGCCT	CTCTGCTCCCGTGGAAATCCATGATCCCTGATAGGCAAGGTT	120
Db	698	CTCACCGAGTCGGTGCCT	CTTGCTCCCGTGGAAATCCATGATCCCTGATAGGCAAGGTT	757
QY	121	TACACAGCAGAAACCCAGTGC	CCAGGCATCGGGCTGGGCAAAATAATTA	180
Db	758	TACACAGCAGAAACCCAGTGC	CCAGGCATCGGGCTGGGCAAAATAATTA	817
QY	181	AATCTGTATGGGATGCGAC	CCCTGTGCGCTGAAGAACCCGAGGTGACGTG	240
Db	818	AATCTGTATGGGATGCGAC	CCCTGTGCGCTGAAGAACCCGAGGTGACGTG	877
QY	241	GAGTACTGTGATGTGCCCT	CTCTGCCACCTCGGGCTGAGACAGTACAGCCAGCCTCAG	300
Db	878	GAGTACTGTGATGTGCCCT	CTCTGCCACCTCGGGCTGAGACAGTACAGCCAGCCTCAG	937
QY	301	TTTCGCATCAAGAGAGGGCT	TTTCGCCGACATCGCCTCCACCCCTGGCAGGCTGCCATC	360
Db	938	TTTCGCATCAAGAGAGGGCT	TTTCGCCGACATCGCCTCCACCCCTGGCAGGCTGCCATC	997
QY	361	TTTGGCAAGCACAGGAGGT	CGCCGGAGAGCGGTTCTGTGCGGGGCACTCATCAGC	420
Db	998	TTTGGCAAGCACAGGAGGT	CGCCGGAGAGCGGTTCTGTGCGGGGCACTCATCAGC	1057
QY	421	TCTGTCTGATTTCTCT	TGCGGCCACTGTCTCCAGGAGAGGTTTCGCGCCCAACCACTG	480
Db	1058	TCTGTCTGATTTCTCT	TGCGGCCCACTGTCTCCAGGAGAGGTTTCGCGCCCAACCACTG	1117
QY	481	ACGGTGTATCTTGGCGAGAA	CATACCGGGTGGTCCCTGGCGAGGAGCAGAAATTTGAA	540
Db	1118	ACGGTGTATCTTGGCGAGAA	CATACCGGGTGGTCCCTGGCGAGGAGCAGAAATTTGAA	1177
QY	541	GTGCAAAATACATTTGTCC	ATAAGGAATTCGATGATGACACTTACGACAATGACATTCGC	600
Db	1178	GTGCAAAATACATTTGTCC	ATAAGGAATTCGATGATGACACTTACGACAATGACATTCGC	1237
QY	601	CTGCTGACGTGAAATCGGA	ATTGCTCCGCTGTGCCAGGAGACAGGTGGTCCGCACT	660
Db	1238	CTGCTGACGTGAAATCGGA	ATTGCTCCGCTGTGCCAGGAGACAGGTGGTCCGCACT	1297
QY	661	GTGTGCTTCCCGGGGGAC	CTGCAGCTCCGAGCTGGAGCTGTGAGCTCTCCGGC	720
Db	1298	GTGTGCTTCCCGGGGGAC	CTGCAGCTCCGAGCTGGAGCTGTGAGCTCTCCGGC	1357
QY	721	TACGGCAAGCATGAGGC	TTTGTCTCTTTCTATTTCGAGCGGCTGAAAGGAGGCTCATGTC	780
Db	1358	TACGGCAAGCATGAGGC	TTTGTCTCTTTCTATTTCGAGCGGCTGAAAGGAGGCTCATGTC	1417
QY	781	AGACTGTACCATCCAGCC	GTGCATCAACATTTACTTAAACAGAACGTACCCGAC	840
Db	1418	AGACTGTACCATCCAGCC	GTGCATCAACATTTACTTAAACAGAACGTACCCGAC	1477
QY	841	AACATGTGTGTCTGAGAC	ACTTCGAGCGGGGGCCCAAGGCAAACTTTGCAGCAGGCC	900





Qy 841 AACATGCTGTGTGGAGACACTCGAGCGGGCCCCAGGCCAAACTTGCACGACGCC 900  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||  
Qy 844 AACATGCTGTGTGGAGACACTCGAGCGGGCCCCAGGCCAAACTTGCACGACGCC 903  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||  
Qy 901 TGCCAGGGCGATTGCGGAGGGCCCCCTGGTGTCTGTAAAGATGGCCGCGATTTGGTG 960  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||  
Qy 904 TGCCAGGGCGATTGCGGAGGGCCCCCTGGTGTCTGTAAAGATGGCCGCGATTTGGTG 963  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||  
Qy 961 GGCATCATCAGCTGGGGCTGGGCTGTGGACAGAGGATGCCCGGGTGTGTACACAAAG 1020  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||  
Qy 964 GGCATCATCAGCTGGGGCTGGGCTGTGGACAGAGGATGCCCGGGTGTGTACACAAAG 1023  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||  
Qy 1021 GTTACCAACTACCTAGACTGGATTCTGTGACACACATGCGACCGTGA 1065  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||  
Qy 1024 GTTACCAACTACCTAGACTGGATTCTGTGACACACATGCGACCGTGA 1068  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||

Search completed: August 19, 2003, 08:35:12  
Job time : 85.3499 secs



Result No.	Score	Query Match	Length	DB	ID	Description
1	1065	100.0	1065	11	US-09-987-455-4	Sequence
2	1065	100.0	1065	11	US-09-987-455-7	Sequence
3	1063.4	99.8	1689	9	US-09-969-271-6	Sequence 6
4	1063.4	99.8	2509	14	US-10-193-656-7	Sequence
5	1063.4	99.8	2519	9	US-09-969-271-5	Sequence 5
6	1062	99.7	1128	11	US-09-987-455-5	Sequence
7	1062	99.7	1128	11	US-09-987-455-5	Sequence
8	1061.8	99.7	2641	10	US-09-974-298-144	Sequence
9	230.2	21.6	329	12	US-10-007-926A-433	Sequence
10	191.6	18.0	1212	10	US-09-880-503-15	Sequence
11	185	17.4	1415	14	US-10-198-846-12748	Sequence
12	170.8	16.0	1475	9	US-09-735-705-122	Sequence 1
13	170.8	16.0	1475	10	US-09-850-716A-122	Sequence
14	170.8	16.0	1475	10	US-09-897-778-122	Sequence
15	170.8	16.0	1475	11	US-09-466-396A-122	Sequence
16	170.8	16.0	1475	12	US-10-117-982-122	Sequence

Db 61 CTCACGAGTGGGTGCTCTCTGCTCCGCTCGGTGAATTCATGATCTCATAGGCAAGTT 120  
Qy 121 TACACAGACAGAACCCAGTGCCTCCAGGACTGGGCTGGGCAACATAATTAAGTCCGG 180  
Db 121 TACACAGACAGAACCCAGTGCCTCCAGGACTGGGCTGGGCAACATAATTAAGTCCGG 180  
Qy 181 AATCTGATGGGATGCCAAGCCCTGGTGCACAGTGTGAAGAACCGCAGGCTGACGTGG 240  
Db 181 AATCTGATGGGATGCCAAGCCCTGGTGCACAGTGTGAAGAACCGCAGGCTGACGTGG 240  
Qy 241 GAGTACTGTGATGTCCTCTCTCTGCTCCACCTGGGCTGGAGACAGTACAGCCAGCTCAG 300  
Db 241 GAGTACTGTGATGTCCTCTCTCTGCTCCACCTGGGCTGGAGACAGTACAGCCAGCTCAG 300  
Qy 301 TTTCGCATCAAGAGGAGGCTTTCGCGACATCGCTCCACCCCTGCGAGGCTGCCATC 360  
Db 301 TTTCGCATCAAGAGGAGGCTTTCGCGACATCGCTCCACCCCTGCGAGGCTGCCATC 360  
Qy 361 TTTCGCAAGCAGGAGGCTGCGCGGAGAGGCTTCTGTCGGGGGATCTCATCAGC 420  
Db 361 TTTCGCAAGCAGGAGGCTGCGCGGAGAGGCTTCTGTCGGGGGATCTCATCAGC 420  
Qy 421 TCCTGCTGGAATTCCTCTGCGGCCACTGCTTCCAGGAGAGGTTTCCGCCCCACCACTG 480  
Db 421 TCCTGCTGGAATTCCTCTGCGGCCACTGCTTCCAGGAGAGGTTTCCGCCCCACCACTG 480  
Qy 481 ACGGTGATCTGGGACAGACATACCGGGTGGTCCCTGCGAGGAGGACAGAAATTTGAA 540  
Db 481 ACGGTGATCTGGGACAGACATACCGGGTGGTCCCTGCGAGGAGGACAGAAATTTGAA 540  
Qy 541 GTCCGAAATAATGATGTCATGAAGAAATTCGATGATGACACTTACGACAAATGCAAG 600  
Db 541 GTCCGAAATAATGATGTCATGAAGAAATTCGATGATGACACTTACGACAAATGCAAG 600  
Qy 601 CTGCTGCAGCTGAATTCGATTCGTCCTTCTATTCGAGCGGCTCAAGGAGGCTCATGTC 780  
Db 601 CTGCTGCAGCTGAATTCGATTCGTCCTTCTATTCGAGCGGCTCAAGGAGGCTCATGTC 780  
Qy 781 AGACTGTATCCCATCCAGCGGCTGCACATCAACATTTTAAACAGAACAGTCAACGAC 840  
Db 781 AGACTGTATCCCATCCAGCGGCTGCACATCAACATTTTAAACAGAACAGTCAACGAC 840  
Qy 841 AACATGCTGTGTGAGAGACTCGAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 900  
Db 841 AACATGCTGTGTGAGAGACTCGAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 900  
Qy 901 TSCCAGGCGGATTCGCGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 960  
Db 901 TSCCAGGCGGATTCGCGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 960  
Qy 961 GGCATCATCAGCTGGGCGGCTGGGAGAGAGGATGTCGCGGCGGCTGTGACAAAG 1020  
Db 961 GGCATCATCAGCTGGGCGGCTGGGAGAGAGGATGTCGCGGCGGCTGTGACAAAG 1020  
Qy 1021 GTTACCACTACCTAGATGGATTCGTGACAAATGCGACCGTGA 1065  
Db 1021 GTTACCACTACCTAGATGGATTCGTGACAAATGCGACCGTGA 1065

RESULT 2  
US-09-987-455-7  
; Sequence 7, Application US/09987455  
; Publication No. US20030049729A1  
; GENERAL INFORMATION:

; APPLICANT: Aranya Manosroi  
; APPLICANT: Jiradej Manosroi  
; APPLICANT: Chatchai Tavapiwatana  
; APPLICANT: Friedrich Goetz  
; APPLICANT: Rolf-Guenther Werner  
; TITLE OF INVENTION: Methods for Large Scale Production of Recombinant  
; TITLE OF INVENTION: DNA-Derived tPA or K2S Molecules  
; FILE REFERENCE: 0652.2190001  
; CURRENT APPLICATION NUMBER: US/09/987,455  
; PRIORITY FILING DATE: 2001-11-14  
; PRIOR APPLICATION NUMBER: 60/268,574  
; PRIOR FILING DATE: 2001-02-15  
; PRIOR APPLICATION NUMBER: GB 0027779.8  
; PRIOR FILING DATE: 2000-11-14  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: PatentIn ver. 2.1  
; SEQ ID NO 7  
; LENGTH: 1065  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: coding  
; OTHER INFORMATION: sequence for K2S protein  
US-09-987-455-7

Query Match 100.0%; Score 1065; DB 11; Length 1065;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1065; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 TCTGAGGAAACAGTGTCTCTCTTTGGAAATGGGTGAGCTACCGTGGCAGCAGCAGC 60  
Db 1 TCTGAGGAAACAGTGTCTCTCTTTGGAAATGGGTGAGCTACCGTGGCAGCAGCAGC 60  
Qy 61 CTACCCAGTGGGTGCTCTCTGCTCCGCTGGAAATTCATGATCTCATAGGCAAGTT 120  
Db 61 CTACCCAGTGGGTGCTCTCTGCTCCGCTGGAAATTCATGATCTCATAGGCAAGTT 120  
Qy 121 TACACAGACAGAACCCAGTGCCTCCAGGACTGGGCTGGGCAACATAATTAAGTCCGG 180  
Db 121 TACACAGACAGAACCCAGTGCCTCCAGGACTGGGCTGGGCAACATAATTAAGTCCGG 180  
Qy 181 AATCTGATGGGATGCCAAGCCCTGGTGCACAGTGTGAAGAACCGCAGGCTGACGTGG 240  
Db 181 AATCTGATGGGATGCCAAGCCCTGGTGCACAGTGTGAAGAACCGCAGGCTGACGTGG 240  
Qy 241 GAGTACTGTGATGTCCTCTCTGCTCCACCTGGGCTGGAGACAGTACAGCCAGCTCAG 300  
Db 241 GAGTACTGTGATGTCCTCTCTGCTCCACCTGGGCTGGAGACAGTACAGCCAGCTCAG 300  
Qy 301 TTTCGCATCAAGAGGAGGCTTTCGCGACATCGCTCCACCCCTGCGAGGCTGCCATC 360  
Db 301 TTTCGCATCAAGAGGAGGCTTTCGCGACATCGCTCCACCCCTGCGAGGCTGCCATC 360  
Qy 361 TTTCGCAAGCAGGAGGCTGCGCGGAGAGGCTTCTGTCGGGGGATCTCATCAGC 420  
Db 361 TTTCGCAAGCAGGAGGCTGCGCGGAGAGGCTTCTGTCGGGGGATCTCATCAGC 420  
Qy 421 TCCTGCTGGAATTCCTCTGCGGCCACTGCTTCCAGGAGAGGTTTCCGCCCCACCACTG 480  
Db 421 TCCTGCTGGAATTCCTCTGCGGCCACTGCTTCCAGGAGAGGTTTCCGCCCCACCACTG 480  
Qy 481 ACGGTGATCTGGGACAGACATACCGGGTGGTCCCTGCGAGGAGGACAGAAATTTGAA 540  
Db 481 ACGGTGATCTGGGACAGACATACCGGGTGGTCCCTGCGAGGAGGACAGAAATTTGAA 540  
Qy 541 GTCCGAAATAATGATGTCATGAAGAAATTCGATGATGACACTTACGACAAATGCAAG 600  
Db 541 GTCCGAAATAATGATGTCATGAAGAAATTCGATGATGACACTTACGACAAATGCAAG 600  
Qy 601 CTGCTGCAGCTGAATTCGATTCGTCCTTCTATTCGAGCGGCTCAAGGAGGCTCGGCACT 660  
Db 601 CTGCTGCAGCTGAATTCGATTCGTCCTTCTATTCGAGCGGCTCAAGGAGGCTCGGCACT 660

```
Qy 661 GTGTGCTTCCCGCGGAGCTCAGCTGCCGGA CTGGA CGGAGTGTGAGCTCTCCGGC 720
Db 661 GTGTGCTTCCCGCGGAGCTCAGCTGCCGGA CTGGA CGGAGTGTGAGCTCTCCGGC 720
Qy 721 TACGGCAAGCATGAGCGCTTGTCTCTTCTTATTCGGAGCGCTGAAGGAGCTCATGTC 780
Db 721 TACGGCAAGCATGAGCGCTTGTCTCTTCTTATTCGGAGCGCTGAAGGAGCTCATGTC 780
Qy 781 AGACTGTACCCATCCAGCGCTGCA CATCAACAATTTTACTTAA CAGAACAGTCAACGAC 840
Db 781 AGACTGTACCCATCCAGCGCTGCA CATCAACAATTTTACTTAA CAGAACAGTCAACGAC 840
Qy 841 AACATGCTGTGTGCGAGACACTCGGAGCGCGCGGCCCGGAGGAGCTTGCACGACGCC 900
Db 841 AACATGCTGTGTGCGAGACACTCGGAGCGCGCGGCCCGGAGGAGCTTGCACGACGCC 900
Qy 901 TGCCAGGCGGATTCGGAGGCGCCCTGTGTGTCTGAACGATGGCGCATGACTTTGGTG 960
Db 901 TGCCAGGCGGATTCGGAGGCGCCCTGTGTGTCTGAACGATGGCGCATGACTTTGGTG 960
Qy 961 GGCATCATCAGTGGGGCTGGGCTGTGGA CAGAGGATGTCCC GGGTGTGTACACAAAG 1020
Db 961 GGCATCATCAGTGGGGCTGGGCTGTGGA CAGAGGATGTCCC GGGTGTGTACACAAAG 1020
Qy 1021 GTTACCAACTACCTAGACTGATTCGTGCA CAACATGCGACCGTGA 1065
Db 1021 GTTACCAACTACCTAGACTGATTCGTGCA CAACATGCGACCGTGA 1065
```

## RESULT 3

```
US-09-969-271-6
; Sequence 6, Application US/09969271
; Patent No. US20020098179A1
; GENERAL INFORMATION:
; APPLICANT: Pfizer Inc. (All designated States except GB and EP (GB));
; APPLICANT: Pfizer Limited (GB and EP (GB) only)
; TITLE OF INVENTION: Pharmaceutical Combinations
; FILE REFERENCE: PCS10951APME
; CURRENT APPLICATION NUMBER: US/09/969,271
; CURRENT FILING DATE: 2001-10-01
; PRIOR APPLICATION NUMBER: GB 0025473.0
; PRIOR FILING DATE: 2000-10-17
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 1689
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-969-271-6
```

```
Query Match 99.8%; Score 1063.4; DB 9; Length 1689;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1064; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
Qy 1 TCTGAGGAAACAGTACTGCTACTTTGGGAATGGGTGAGCTACCGTGGCAACGACAGC 60
Db 625 TCTGAGGAAACAGTACTGCTACTTTGGGAATGGGTGAGCTACCGTGGCAACGACAGC 684
Qy 61 CTCACCGAGTCGGGTGCTCTGCTCCCGTGGAAATCCATGATCCCTGATAGGCAAGTT 120
Db 685 CTCACCGAGTCGGGTGCTCTGCTCCCGTGGAAATCCATGATCCCTGATAGGCAAGTT 744
Qy 121 TACACAGACAGAAACCCAGTGCACGAGCACTGGGCTGGGCAACATAATTAATCTGCGG 180
Db 745 TACACAGACAGAAACCCAGTGCACGAGCACTGGGCTGGGCAACATAATTAATCTGCGG 804
Qy 181 AATCCTGATGGGATGCCAACCCCTGGTGCACAGTGTGAAGAACCGGAGGCTGACGTGG 240
Db 805 AATCCTGATGGGATGCCAACCCCTGGTGCACAGTGTGAAGAACCGGAGGCTGACGTGG 864
Qy 241 GAGTACTGTGATGTGCGCTCTGCTCCACCTGCGGCTGAGACATGACGACCCCTCAG 300
Db 865 GAGTACTGTGATGTGCGCTCTGCTCCACCTGCGGCTGAGACATGACGACCCCTCAG 924
```

```
Qy 301 TTTCCGATCAAAAGAGAGGCTCTTCCCGACATCGCTCCACCCCTGCGAGGCTGCCATC 360
Db 925 TTTCCGATCAAAAGAGAGGCTCTTCCCGACATCGCTCCACCCCTGCGAGGCTGCCATC 984
Qy 361 TTTCCGACACACAGAGGTGCGCCGGAGAGCGGTTCCTGTGCGGGGGCATACTCATCAGC 420
Db 985 TTTCCGACACACAGAGGTGCGCCGGAGAGCGGTTCCTGTGCGGGGGCATACTCATCAGC 1044
Qy 421 TCCTGCTGGATTCTCTGCGCGCCACTGCTTCCAGGAGAGGTTTCGCGCCCAACACCTG 480
Db 1045 TCCTGCTGGATTCTCTGCGCGCCACTGCTTCCAGGAGAGGTTTCGCGCCCAACACCTG 1104
Qy 481 ACGTGTATCTTTGGCGACAAACATACCGGTGCTCCCTGGCGAGGAGGACGAAATTTGAA 540
Db 1105 ACGTGTATCTTTGGCGACAAACATACCGGTGCTCCCTGGCGAGGAGGACGAAATTTGAA 1164
Qy 541 GTCCAAAATACATGTTGTCATTAAGGAATTCGATGATGACACTTACGACAAATGACATGCG 600
Db 1165 GTCCAAAATACATGTTGTCATTAAGGAATTCGATGATGACACTTACGACAAATGACATGCG 1224
Qy 601 CTGCTGCACTGAAATCGGATTTCGTCGCGCTGTGCCAGAGAGAGCGTGGTCCGCACT 660
Db 1225 CTGCTGCACTGAAATCGGATTTCGTCGCGCTGTGCCAGAGAGAGCGTGGTCCGCACT 1284
Qy 661 GTGTGCTTCCCGCGGAGCTTCGAGCTGCCGACTGGACGGAGTGTGAGCTCTCCGGC 720
Db 1285 GTGTGCTTCCCGCGGAGCTTCGAGCTGCCGACTGGACGGAGTGTGAGCTCTCCGGC 1344
Qy 721 TACGGCAAGCATGAGGCTTGTCTCTTCTTATTCGGAGCGGCTGAAGGAGGCTCATGTC 780
Db 1345 TACGGCAAGCATGAGGCTTGTCTCTTCTTATTCGGAGCGGCTGAAGGAGGCTCATGTC 1404
Qy 781 AGACTGTACCCATCCAGCGCTGCACATCAACAATTTACTTAA CAGAACAGTCAACGAC 840
Db 1405 AGACTGTACCCATCCAGCGCTGCACATCAACAATTTACTTAA CAGAACAGTCAACGAC 1464
Qy 841 AACATGCTGTGTGCTGAGACACTCGGAGCGCGGGCCCGAGGCAAACTTTGACGAGGCC 900
Db 1465 AACATGCTGTGTGCTGAGACACTCGGAGCGCGGGCCCGAGGCAAACTTTGACGAGGCC 1524
Qy 901 TGCAGGCGGATTCGGAGGAGCGCCCTGCTGTGTCTGAACGATGGCGCGCATGACTTTGGTG 960
Db 1525 TGCAGGCGGATTCGGAGGAGCGCCCTGCTGTGTCTGAACGATGGCGCGCATGACTTTGGTG 1584
Qy 961 GGCATCATCAGTGGGCGCTGGGCTGTGGA CAGAGGATGTGCCAGAGGATGTCCCGGGTGTGTACCAAG 1020
Db 1585 GGCATCATCAGTGGGCGCTGGGCTGTGGA CAGAGGATGTGCCAGAGGATGTCCCGGGTGTGTACCAAG 1065
Qy 1021 GTTACCAACTACCTAGACTGATTCGTGCA CAACATGCGACCGTGA 1065
Db 1645 GTTACCAACTACCTAGACTGATTCGTGCA CAACATGCGACCGTGA 1689
```

## RESULT 4

```
US-10-193-656-7
; Sequence 7, Application US/10193656
; Publication No. US20030096733A1
; GENERAL INFORMATION:
; APPLICANT: NY, Tor
; APPLICANT: HOLMDAHL, Rikard
; APPLICANT: Li, Jinan
; TITLE OF INVENTION: NOVEL DRUG TARGETS FOR ARTHRITIS
; FILE REFERENCE: 3810/1J577-US3
; CURRENT APPLICATION NUMBER: US/10/193,656
; CURRENT FILING DATE: 2002-07-10
; PRIOR APPLICATION NUMBER: US 60/304,461
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US 60/304,490
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US 60/305,182
; PRIOR FILING DATE: 2001-07-13
; NUMBER OF SEQ ID NOS: 18
```

```
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
; LENGTH: 2509
; TYPE: DNA
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: GenBank / X07393
; DATABASE ENTRY DATE: 1995-03-27
; RELEVANT RESIDUES: (1)..(2509)
US-10-193-656-7

Query Match      99.8%; Score 1063.4; DB 14; Length 2509;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1064; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TCTGAGGGAACAGTACTGCTACTTTGGGAATGGGTGAGCTACCGTGGGACGACGAGC 60
DB 678 TCTGAGGGAACAGTACTGCTACTTTGGGAATGGGTGAGCTACCGTGGGACGACGAGC 737
QY 61 CTCACCGAGTCGGGTGCTCTGCTCCGCTGGGAATCCATGATCCCTGATAGGCAAGTT 120
DB 738 CTCACCGAGTCGGGTGCTCTGCTCCGCTGGGAATCCATGATCCCTGATAGGCAAGTT 797
QY 121 TACACGACAGAAACCCAGTCCGAGCACTGGGCTGGGCAACATTAATCTGCGG 180
DB 798 TACACGACAGAAACCCAGTCCGAGCACTGGGCTGGGCAACATTAATCTGCGG 857
QY 181 AATCTGATGGGATGCCAAGCCCTGTGTCACAGTGTGAAGAACCGGAGCTGACGTG 240
DB 858 AATCTGATGGGATGCCAAGCCCTGTGTCACAGTGTGAAGAACCGGAGCTGACGTG 917
QY 241 GAGTACTGTGATGTCCTCTGCTCCACCTCGGCTGACACAGTACAGCCAGCTCAG 300
DB 918 GAGTACTGTGATGTCCTCTGCTCCACCTCGGCTGACACAGTACAGCCAGCTCAG 977
QY 301 TTTGCGATCAAGAGGGCTCTTTCGCGACATCGCTCCACCCCTGGCAGGCTGCCATC 360
DB 978 TTTGCGATCAAGAGGGCTCTTTCGCGACATCGCTCCACCCCTGGCAGGCTGCCATC 1037
QY 361 TTTGCGAAGCAAGAGTCCGCGAGAGCGGTTCTGTGCGGGGGATATCTATCAGC 420
DB 1038 TTTGCGAAGCAAGAGTCCGCGAGAGCGGTTCTGTGCGGGGGATATCTATCAGC 1097
QY 421 TCCTGCTGATTTCTCTGCGCGCCACTGCTTCAGAGAGAGTTTCGCCCCACCACTG 480
DB 1098 TCCTGCTGATTTCTCTGCGCGCCACTGCTTCAGAGAGAGTTTCGCCCCACCACTG 1157
QY 481 ACGGTGATCTTGGGAGAACATACCGGGTGGTCCCTGGCGAGGAGGACGAAATTTGAA 540
DB 1158 ACGGTGATCTTGGGAGAACATACCGGGTGGTCCCTGGCGAGGAGGACGAAATTTGAA 1217
QY 541 GTCGAAATATATGTCATTAAGGAATTCATGATGACATTCAGCAATGACATTCGCG 600
DB 1218 GTCGAAATATATGTCATTAAGGAATTCATGATGACATTCAGCAATGACATTCGCG 1277
QY 601 CTGCTGACGCTGAATTCGCTCCGCTGTCGCCAGGAGAGCGGTGGTCCGCACT 660
DB 1278 CTGCTGACGCTGAATTCGCTCCGCTGTCGCCAGGAGAGCGGTGGTCCGCACT 1337
QY 661 GTGTGCTTTCCCGCGGACCTGACGCTGCGGACCTGGAGGAGTGTGAGCTCTCCGGC 720
DB 1338 GTGTGCTTTCCCGCGGACCTGACGCTGCGGACCTGGAGGAGTGTGAGCTCTCCGGC 1397
QY 721 TACGCAAGCATAGGCGCTTGTCTCTTTATTCGAGCGGCTGAGGAGGCTCATGTC 780
DB 1398 TACGCAAGCATAGGCGCTTGTCTCTTTATTCGAGCGGCTGAGGAGGCTCATGTC 1457
QY 781 AGACTGTACCATCCAGCGCTGACATCAACAATTTACTTAACAGAAAGTACCGGAC 840
DB 1458 AGACTGTACCATCCAGCGCTGACATCAACAATTTACTTAACAGAAAGTACCGGAC 1517
QY 841 AACATGCTGTGTGTGGAGACATCGGAGCGCGGGCCCGAGGCAAACTTTCACGAGCC 900
DB 1518 AACATGCTGTGTGTGGAGACATCGGAGCGCGGGCCCGAGGCAAACTTTCACGAGCC 1577
QY 901 TGCAGGGCGAATTCGGAGAGCCCTCTGCTGTCTGACCATGCGCGCATGACTTTGGTG 960
DB 1578 TGCAGGGCGAATTCGGAGAGCCCTCTGCTGTCTGACCATGCGCGCATGACTTTGGTG 1637
QY 961 GGCAATCATAGCTGGGGCTTGGGCTGTGGACAGAGGATGTCCTGGGTGTACACAAAG 1020
DB 1638 GGCAATCATAGCTGGGGCTTGGGCTGTGGACAGAGGATGTCCTGGGTGTGTACACAAAG 1697
QY 1021 GTTACCAACTACCTAGACTGGATTCTGTGACAAACATGCGACCGTGA 1065
DB 1698 GTTACCAACTACCTAGACTGGATTCTGTGACAAACATGCGACCGTGA 1742

RESULT 5
US-09-969-271-5
; Sequence 5, Application US/09969271
; Patent No. US20020098179A1
; GENERAL INFORMATION:
; APPLICANT: Pfizer Inc. (All designated States except GB and EP (GB));
; APPLICANT: Pfizer Limited (GB and EP (GB) only)
; TITLE OF INVENTION: Pharmaceutical Combinations
; FILE REFERENCE: PCS10951APME
; CURRENT APPLICATION NUMBER: US/09/969,271
; CURRENT FILING DATE: 2001-10-01
; PRIOR APPLICATION NUMBER: GB 0025473.0
; PRIOR FILING DATE: 2000-10-17
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 2519
; TYPE: DNA
; ORGANISM: Homo sapiens
; ORGANISM: Homo sapiens
US-09-969-271-5

Query Match      99.8%; Score 1063.4; DB 9; Length 2519;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1064; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TCTGAGGGAACAGTACTGCTACTTTGGGAATGGGTGAGCTACCGTGGGACGACGAGC 60
DB 701 TCTGAGGGAACAGTACTGCTACTTTGGGAATGGGTGAGCTACCGTGGGACGACGAGC 760
QY 61 CTACCGAGTCGGGTGCTCTGCTCCGCTGGGAATTCCTGATGATGAGGCAAGTT 120
DB 761 CTACCGAGTCGGGTGCTCTGCTCCGCTGGGAATTCCTGATGATGAGGCAAGTT 820
QY 121 TACACAGCACAGAACCCAGTGGCCAGGCACTGGGCTGGGCAACATTAATCTGCGG 180
DB 821 TACACAGCACAGAACCCAGTGGCCAGGCACTGGGCTGGGCAACATTAATCTGCGG 880
QY 181 AATCTGATGGGATGCCAAGCCCTGTGTGCGAGTGTGAAGAACCCGAGGCTGACGTG 240
DB 881 AATCTGATGGGATGCCAAGCCCTGTGTGCGAGTGTGAAGAACCCGAGGCTGACGTG 940
QY 241 GAGTACTGTGATGTCCTCTGCTCCACTGCGGCTGAGACAGTACAGCCAGCTCAG 300
DB 941 GAGTACTGTGATGTCCTCTGCTCCACTGCGGCTGAGACAGTACAGCCAGCTCAG 1000
QY 301 TTTGCGATCAAGAGGAGGCTCTTTCGCGACATCGCTCCACCCCTGGCAGGCTGCCATC 360
DB 1001 TTTGCGATCAAGAGGAGGCTCTTTCGCGACATCGCTCCACCCCTGGCAGGCTGCCATC 1060
QY 361 TTTGCGAAGCAAGAGTCCGCGAGAGCGGTTCTGTGCGGGGGATATCTATCAGC 420
DB 1061 TTTGCGAAGCAAGAGTCCGCGAGAGCGGTTCTGTGCGGGGGATATCTATCAGC 1120
QY 421 TCCTGCTGATTTCTCTGCGCGCCACTGCTTCAGGAGAGGTTTTCGCCCCACCACTG 480
DB 1121 TCCTGCTGATTTCTCTGCGCGCCACTGCTTCAGGAGAGGTTTTCGCCCCACCACTG 1180
QY 481 ACGGTGATCTTGGGAGAACATACCGGGTGGTTCCTGCGGAGGAGGAGGAAATTTGAA 540
```

Db 1181 ACAGGATCTTTGGGAGAACATACCGGGTGGTCCCTGGCGAGGAGCAGCAAAATTTGAA 1240  
Qy 541 GTGCAAAATACATTTGTCATAAGGAATTCGATGATGACATTCAGCAATGACATTCGG 600  
Db 1241 GTGCAAAATACATTTGTCATAAGGAATTCGATGATGACATTCAGCAATGACATTCGG 1300  
Qy 601 CTGTGACGTGAATTCGATTCGTCGCTGTGTCAGGAGAGCAGTGGTCCGCACT 660  
Db 1301 CTGTGACGTGAATTCGATTCGTCGCTGTGTCAGGAGAGCAGTGGTCCGCACT 1360  
Qy 661 GTGTGCTTCCCGCGGAGCTGAGCTGCGGAGCTGCGAGCTGCGAGCTGCGAGCTTCCGGC 720  
Db 1361 GTGTGCTTCCCGCGGAGCTGAGCTGCGGAGCTGCGAGCTGCGAGCTTCCGGC 1420  
Qy 721 TACGCAAGCATAGGCGCTTGTCTCTTCTTATTCGAGCGGCTGAAGGAGGCTCATGTC 780  
Db 1421 TACGCAAGCATAGGCGCTTGTCTCTTCTTATTCGAGCGGCTGAAGGAGGCTCATGTC 1480  
Qy 781 AGATGTACCATCCAGCCGCTGCACATCAACAATTTTAAACAGAACAGTCAACGAC 840  
Db 1481 AGATGTACCATCCAGCCGCTGCACATCAACAATTTTAAACAGAACAGTCAACGAC 1540  
Qy 841 AACATGCTGTGCTGGAGACACTCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 900  
Db 1541 AACATGCTGTGCTGGAGACACTCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1600  
Qy 901 TGCCAGGCGGATTCGGGAGGCGGCGGCTGTGCTGTAACGATGGCGGCGGCGGCGGCGGCGG 960  
Db 1601 TGCCAGGCGGATTCGGGAGGCGGCGGCTGTGCTGTAACGATGGCGGCGGCGGCGGCGGCGG 1660  
Qy 961 GGCAATCATCAGTGGGCGGCTGGGCTGTGGAGAGAGTGTCCGGGTGTGTACACAAAG 1020  
Db 1661 GGCAATCATCAGTGGGCGGCTGGGCTGTGGAGAGAGTGTCCGGGTGTGTACACAAAG 1720  
Qy 1021 GTTACCAACTACCTAGATCGATTCGTGACAAACATCGACCGTGA 1065  
Db 1721 GTTACCAACTACCTAGATCGATTCGTGACAAACATCGACCGTGA 1765

RESULT 6

US-09-987-455-2  
; Sequence 2, Application US/09987455  
; Publication No. US20030049729A1  
; GENERAL INFORMATION:  
; APPLICANT: Aranya Manosroi  
; APPLICANT: Jiradej Manosroi  
; APPLICANT: Chatchai Tayapiwatana  
; APPLICANT: Friedrich Goetz  
; APPLICANT: Rolf-Guenther Werner  
; TITLE OF INVENTION: Methods for Large Scale Production of Recombinant  
; FILE REFERENCE: 0652.2190001  
; CURRENT APPLICATION NUMBER: US/09/987,455  
; CURRENT FILING DATE: 2001-11-14  
; PRIOR APPLICATION NUMBER: 60/268,574  
; PRIOR FILING DATE: 2001-02-15  
; PRIOR APPLICATION NUMBER: GB 0027779.8  
; PRIOR FILING DATE: 2000-11-14  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 1128  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURES:  
; OTHER INFORMATION: Description of Artificial Sequence: coding  
; OTHER INFORMATION: sequence for OmpA-K2S fusion protein  
US-09-987-455-2

Query Match 99.7%; Score 1062; DB 11; Length 1128;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1062; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCTGAGGAAACAGTGACTGCTACTTTTGGGAATGGGTACGCTACCGTGGCAGCGCACAGC 60  
Db 67 TCTGAGGAAACAGTGACTGCTACTTTTGGGAATGGGTACGCTACCGTGGCAGCGCACAGC 126  
Qy 61 CTACCCAGTGGGTGCTCTCTGCTCCGCTGGGAATTCATGATCCTGATAGGCAAGTT 120  
Db 127 CTACCCAGTGGGTGCTCTCTGCTCCGCTGGGAATTCATGATCCTGATAGGCAAGTT 186  
Qy 121 TACACAGCAGAAACCCAGTCCAGGCACTGGGCTGGGCAAAACAATAATTTACTGCCGG 180  
Db 187 TACACAGCAGAAACCCAGTCCAGGCACTGGGCTGGGCAAAACAATAATTTACTGCCGG 246  
Qy 181 AATCTGATGGGATGCCAAGCCCTGTGTGCCACGCTGTCTGTAAGAAACCGCAGGCTGACGTGG 240  
Db 247 AATCTGATGGGATGCCAAGCCCTGTGTGCCACGCTGTCTGTAAGAAACCGCAGGCTGACGTGG 306  
Qy 241 GAGTACTGTGATGTGCTCTCTGCTCCAGCTGCGGCTGAGACAGTACAGCAGGCTCAG 300  
Db 307 GAGTACTGTGATGTGCTCTCTGCTCCAGCTGCGGCTGAGACAGTACAGCAGGCTCAG 366  
Qy 301 TTTCGCAATCAAGAGAGGCTCTTCCGCGACATCGCTCCACCCCTGGCAGGCTGCCATC 360  
Db 367 TTTCGCAATCAAGAGAGGCTCTTCCGCGACATCGCTCCACCCCTGGCAGGCTGCCATC 426  
Qy 361 TTTCGCAATCAAGAGAGGCTCTTCCGCGAGAGCGGTTCTGTGCGGGGGCATACTCATCAGC 420  
Db 427 TTTCGCAATCAAGAGAGGCTCTTCCGCGAGAGCGGTTCTGTGCGGGGGCATACTCATCAGC 486  
Qy 421 TCCTGTGGAATTCCTCTGCGGCGGCTGCTTCAGGAGAGGTTTCCGCGCCACCACTG 480  
Db 487 TCCTGTGGAATTCCTCTGCGGCGGCTGCTTCAGGAGAGGTTTCCGCGCCACCACTG 546  
Qy 481 ACAGTGTCTTGGGCAAGACATACCGGCTGGTCCCTGGCGAGGAGGAGCAAAATTTGAA 540  
Db 547 ACAGTGTCTTGGGCAAGACATACCGGCTGGTCCCTGGCGAGGAGGAGCAAAATTTGAA 606  
Qy 541 GTCCAAAAATACATTTGTCATTAAGGAATTCGATGATGACACTTACGCAATGACATTCGCG 600  
Db 607 GTCCAAAAATACATTTGTCATTAAGGAATTCGATGATGACACTTACGCAATGACATTCGCG 666  
Qy 601 CTGTGACGTGAATTCGATTCGCTCTCTTCTTATTCGAGCGGCTGAAGGAGGCTCATGTC 660  
Db 667 CTGTGACGTGAATTCGATTCGCTCTCTTCTTATTCGAGCGGCTGAAGGAGGCTCATGTC 726  
Qy 661 GTGTGCTTCCCGCGGAGCTGTCAGCTGCGGAGCTGCGAGTGTGAGCTCTCCGGC 720  
Db 727 GTGTGCTTCCCGCGGAGCTGTCAGCTGCGGAGCTGCGAGGAGTGTGAGCTCTCCGGC 786  
Qy 721 TACGCAAGCATAGGCGCTTGTCTCTTCTTATTCGAGCGGCTGAAGGAGGCTCATGTC 780  
Db 787 TACGCAAGCATAGGCGCTTGTCTCTTCTTATTCGAGCGGCTGAAGGAGGCTCATGTC 846  
Qy 781 AGACTGTACCATCCAGCCGCTGCACATCAACAATTTTAAACAGAACAGTCAACGAC 840  
Db 847 AGACTGTACCATCCAGCCGCTGCACATCAACAATTTTAAACAGAACAGTCAACGAC 906  
Qy 841 AACATGCTGTGCTGGAGACACTCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 900  
Db 907 AACATGCTGTGCTGGAGACACTCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 966  
Qy 901 TGCCAGGCGGATTCGGGAGGCGGCGGCTGTGTCTGTAACGATGGCGGCGGCGGCGGCGGCGG 960  
Db 967 TGCCAGGCGGATTCGGGAGGCGGCGGCTGTGTCTGTAACGATGGCGGCGGCGGCGGCGGCGG 1026  
Qy 961 GGCAATCATCAGTGGGCGCTGGGCTGTGGAAGAGATGTCCGGGTGTGTACACAAAG 1020  
Db 1027 GGCAATCATCAGTGGGCGCTGGGCTGTGGAAGAGATGTCCGGGTGTGTACACAAAG 1086  
Qy 1021 GTTACCAACTACCTAGATTCGTGACAAACATCGACCG 1062  
Db 1087 GTTACCAACTACCTAGATTCGTGACAAACATCGACCG 1128

RESULT 7  
 US-09-987-455-5  
 ; Sequence 5, Application US/09987455  
 ; Publication No. US20030049729A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Aranya Manosroi  
 ; APPLICANT: Jiradej Manosroi  
 ; APPLICANT: Chatchai Tayapiwatana  
 ; APPLICANT: Friedrich Goetz  
 ; APPLICANT: Rolf-Guenther Werner  
 ; TITLE OF INVENTION: Methods for Large Scale Production of Recombinant  
 ; TITLE OF INVENTION: DNA-Derived tPA or K2S Molecules  
 ; FILE REFERENCE: 0652.2190001  
 ; CURRENT APPLICATION NUMBER: US/09/987,455  
 ; CURRENT FILING DATE: 2001-11-14  
 ; PRIOR APPLICATION NUMBER: 60/268,574  
 ; PRIOR FILING DATE: 2001-02-15  
 ; PRIOR APPLICATION NUMBER: GB 0027779.8  
 ; PRIOR FILING DATE: 2000-11-14  
 ; NUMBER OF SEQ ID NOS: 25  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 5  
 ; LENGTH: 1128  
 ; TYPE: DNA  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: Description of Artificial Sequence: coding  
 ; OTHER INFORMATION: sequence for OmpA-K2S fusion protein  
 US-09-987-455-5  
  
 Query Match 99.7%; Score 1062; DB 11; Length 1128;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1062; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
 QY 1 TCTGAGGAAACAGTACTGCTACTTTGGGAATGGTTCAGCTACCGTGGCACGACAGC 60  
 DB 67 TCTGAGGAAACAGTACTGCTACTTTGGGAATGGTTCAGCTACCGTGGCACGACAGC 126  
 QY 61 CTCACGAGTGGGTGCTCTCTGCTCCGCTGGGAATTCATGATTCATAGGCAAGTT 120  
 DB 127 CTCACGAGTGGGTGCTCTCTGCTCCGCTGGGAATTCATGATTCATAGGCAAGTT 186  
 QY 121 TACACGACAGAAACCCAGTGCACGACCTGGGCTGGGCAACATAATTAATCTGCGG 180  
 DB 187 TACACGACAGAAACCCAGTGCACGACCTGGGCTGGGCAACATAATTAATCTGCGG 246  
 QY 181 AATCTGATGGGATGCAAGCCCTGTGTGCAACGCTGTGCAACGCTGCAAGCTGCA 240  
 DB 247 AATCTGATGGGATGCAAGCCCTGTGTGCAACGCTGTGCAACGCTGCAAGCTGCA 306  
 QY 241 GAGTACTGTGATGCTGCTCTCTGCTCCACCTGCGGCTGACACAGTACAGCCAGCTCAG 300  
 DB 307 GAGTACTGTGATGCTGCTCTCTGCTCCACCTGCGGCTGACACAGTACAGCCAGCTCAG 366  
 QY 301 TTTGCAATCAAGAGGAGGCTCTTCCGCGACATCGCTCCACCTGCGGCTGCGGCTGCGCATC 360  
 DB 367 TTTGCAATCAAGAGGAGGCTCTTCCGCGACATCGCTCCACCTGCGGCTGCGGCTGCGCATC 426  
 QY 361 TTTGCAAGCAAGGAGTGCCTGCGGAGAGCGGTTCTGTGCGGCGGATCACTCATCAGC 420  
 DB 427 TTTGCAAGCAAGGAGTGCCTGCGGAGAGCGGTTCTGTGCGGCGGATCACTCATCAGC 486  
 QY 421 TCCTGCTGATCTCTCTGCGGCGGCTGCTTCCAGGAGAGGTTTCGCGGCGGATCACTCATCAG 480  
 DB 487 TCCTGCTGATCTCTCTGCGGCGGCTGCTTCCAGGAGAGGTTTCGCGGCGGATCACTCATCAG 546  
 QY 481 ACGGTGATCTTTGGGAGAGCAATACCGGCTGTGTCTGCGGAGGAGGAGCAAAATTTGAA 540  
 DB 547 ACGGTGATCTTTGGGAGAGCAATACCGGCTGTGTCTGCGGAGGAGGAGCAAAATTTGAA 606  
 QY 541 GTCGAAAATACATTTGTCATTAAGAAATTCGATGATGACATTAACGAAATGACATTTGCG 600

DB 607 GTCGAAAAATACATTTGTCATTAAGAAATTCGATGATGACACTTACGACAAATGACATTTGCG 666  
 QY 601 CTGCTGCAGCTGAATCGGATTCGTCGCGCTGTCGCCAGAGAGCAGCGTGGTCCGCACT 660  
 DB 667 CTGCTGCAGCTGAATCGGATTCGTCGCGCTGTCGCCAGAGAGCAGCGTGGTCCGCACT 726  
 QY 661 GTGTGCCCTTCCCGCGGAGCCTGTCAGCTGTCGCGGCTGTCGCGGAGTGTGAGCTCTCCGCGC 720  
 DB 727 GTGTGCCCTTCCCGCGGAGCCTGTCAGCTGTCGCGGCTGTCGCGGAGTGTGAGCTCTCCGCGC 786  
 QY 721 TACGGCAAGCATGAGGCTTGTCTCTCTTTCTATTCGAGCGGCTGAAGAGGCTCATGTC 780  
 DB 787 TACGGCAAGCATGAGGCTTGTCTCTCTTTCTATTCGAGCGGCTGAAGAGGCTCATGTC 846  
 QY 781 AGACTGTACCCATCCAGCGCTGCACATCAACAATTTACTTAACAGAAAGTCAACGAC 840  
 DB 847 AGACTGTACCCATCCAGCGCTGCACATCAACAATTTACTTAACAGAAAGTCAACGAC 906  
 QY 841 AACATGCTGTGCTGGAGACACTTCGAGCGGCGGCGGCCAGGCAAACTTTCAGCAGCGCC 900  
 DB 907 AACATGCTGTGCTGGAGACACTTCGAGCGGCGGCGGCCAGGCAAACTTTCAGCAGCGCC 966  
 QY 901 TGCAGGCGGATTCGAGGCGGCTGCTGCTGCTGAACGATGCGCGGCTGCTGCTGCTG 960  
 DB 967 TGCAGGCGGATTCGAGGCGGCTGCTGCTGCTGAACGATGCGCGGCTGCTGCTGCTG 1026  
 QY 961 GGCAATCATCAGCTGGGCGCTGGGCTGGGAGAGGATGTCGCGGCTGCTGACAAAG 1020  
 DB 1027 GGCAATCATCAGCTGGGCGCTGGGCTGGGAGAGGATGTCGCGGCTGCTGACAAAG 1086  
 QY 1021 GTTACCAACTACCTAGACTGGATTCGTGACAAACATGCGACCG 1062  
 DB 1087 GTTACCAACTACCTAGACTGGATTCGTGACAAACATGCGACCG 1128

RESULT 8  
 US-09-974-298-144  
 ; Sequence 144, Application US/09974298  
 ; Patent No. US20020156263A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Chen, Hwei-Mei  
 ; TITLE OF INVENTION: GENES EXPRESSED IN BREAST CANCER  
 ; FILE REFERENCE: PA-0037 P  
 ; CURRENT APPLICATION NUMBER: US/09/974,298  
 ; CURRENT FILING DATE: 2001-10-04  
 ; PRIOR APPLICATION NUMBER: 60/238,331  
 ; PRIOR FILING DATE: 2000-05-10  
 ; NUMBER OF SEQ ID NOS: 194  
 ; SOFTWARE: PERL Program  
 ; SEQ ID NO 144  
 ; LENGTH: 2641  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: misc feature  
 ; OTHER INFORMATION: Incyte ID No. US20020156263A1 1001470CB1  
 ; NAME/KEY: unsure  
 ; LOCATION: 2635  
 ; OTHER INFORMATION: a, t, c, g, or other  
 US-09-974-298-144  
  
 Query Match 99.7%; Score 1061.8; DB 10; Length 2641;  
 Best Local Similarity 99.8%; Pred. No. 0;  
 Matches 1063; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
  
 QY 1 TCTGAGGAAACAGTACTGCTACTTTGGGAATGGGTGAGCTACCGTGGCACGACAGC 60  
 DB 821 TCTGAGGAAACAGTACTGCTACTTTGGGAATGGGTGAGCTACCGTGGCACGACAGC 880  
 QY 61 CTACCCAGTGGGTGCTCTGCTCCGCTGGGAATTCATGATTCCTGATAGGCAAGTT 120  
 DB 881 CTACCCAGTGGGTGCTCTGCTCCGCTGGGAATTCATGATTCCTGATAGGCAAGTT 940

Qy	121	TACACAGCAGAAACCCAGCTGCGCCAGGCACTGGGCTTGGGCCAAACATAATTA	CTGCGG	180
Db	941			1000
		TACACAGCAGAAACCCAGCTGCGCCAGGCACTGGGCTTGGGCCAAACATAATTA	CTGCGG	
Qy	181	AATCCTGATGGGATGCAAGCCCTGTGTGCCAGCTGTGAAGAACCCAGGCTGACGTGG	240	
Db	1001			1060
		AATCCTGATGGGATGCAAGCCCTGTGTGCCAGCTGTGAAGAACCCAGGCTGACGTGG		
Qy	241	GAGTACTGTGATGTGCCCTCTGTCTCCACCTCGCGGCTTGAGACAGTACAGCCAGCCTCAG	300	
Db	1061			1120
		GAGTACTGTGATGTGCCCTCTGTCTCCACCTCGCGGCTTGAGACAGTACAGCCAGCCTCAG		
Qy	301	TTTTCGCATCAAGAGAGGGCTCTTCGCCGCAATGCGCTCCACCCCTGGCAGGTGCCATC	360	
Db	1121			1180
		TTTTCGCATCAAGAGAGGGCTCTTCGCCGCAATGCGCTCCACCCCTGGCAGGTGCCATC		
Qy	361	TTTTCGCAAGCACAGAGAGTCCCGGAGAGCGGTTCCTGTGCGGGGCGATACATCAGC	420	
Db	1181			1240
		TTTTCGCAAGCACAGAGAGTCCCGGAGAGCGGTTCCTGTGCGGGGCGATACATCAGC		
Qy	421	TCCTGTGGAATTCTCTCGCGGCCACTGCTTCCAGGAGAGTTTCGCGCCCAACCACTG	480	
Db	1241			1300
		TCCTGTGGAATTCTCTCTCGCGGCCACTGCTTCCAGGAGAGTTTCGCGCCCAACCACTG		
Qy	481	ACGGTGAATCTTGGCAGAAATACCCGGGTGTCCTGGCCAGGAGGAGAGAGAAATTTGAA	540	
Db	1301			1360
		ACGGTGAATCTTGGCAGAAATACCCGGGTGTCCTGGCCAGGAGGAGAGAGAAATTTGAA		
Qy	541	GTGCAAAAATACATTTGTCCATTAAGGAATTCGATGATGACACTTACGACATGACATTCGG	600	
Db	1361			1420
		GTGCAAAAATACATTTGTCCATTAAGGAATTCGATGATGACACTTACGACATGACATTCGG		
Qy	601	CTGCTGCAGCTGAAATCEGAATTGTCCTCCGCTGTGCCAGGAGACGCTGGTCCGCACT	660	
Db	1421			1480
		CTGCTGCAGCTGAAATCGGAATCGTCCGCTGTGCCAGGAGACGCTGGTCCGCACT		
Qy	661	GTGTGCTTCCCGGGGGACCTGCGAGCTGCCGAGCTGGAAGTGTGAGCTCTCCGGC	720	
Db	1481			1540
		GTGTGCTTCCCGGGGGACCTGCGAGCTGCCGAGCTGGAAGTGTGAGCTCTCCGGC		
Qy	721	TACGGCAAGCATGAGGCTTGTCTCTTTCTATTCCGAGCGCTGAAGGAGGCTCATGTC	780	
Db	1541			1600
		TACGGCAAGCATGAGGCTTGTCTCTTTCTATTCCGAGCGCTGAAGGAGGCTCATGTC		
Qy	781	AGACTGTACCCATCCAGCCGCTGCATCAACAATTTACTTAACGAAACAGTCACCGAC	840	
Db	1601			1660
		AGACTGTACCCATCCAGCCGCTGCATCAACAATTTACTTAACGAAACAGTCACCGAC		
Qy	841	AACATGCTGTGTGTGAGAGACATCTCGAGCGGGGGGCCCCAGGGAATTTGACGAGCC	900	
Db	1661			1720
		AACATGCTGTGTGTGAGAGACATCTCGAGCGGGGGGCCCCAGGGAATTTGACGAGCC		
Qy	901	TGCCAGGGCGATTTGGGAGGCCCTTGTGTGTCGAACGATGCGCGCATGACTTTGGTG	960	
Db	1721			1780
		TGCCAGGGCGATTTGGGAGGCCCTTGTGTGTCGAACGATGCGCGCATGACTTTGGTG		
Qy	961	GGCATCATAGCTGGGGCTTGGGCTGTGGACAGAGGATGTCCTCGGGGTGTGTACAAAG	1020	
Db	1781			1840
		GGCATCATAGCTGGGGCTTGGGCTGTGGACAGAGGATGTCCTCGGGGTGTGTACAAAG		
Qy	1021	GTTTACCACTTACCTAGACTGGATCTGTGTGACAAATGCGACCGTGA	1065	
Db	1841	GTTTACCACTTACCTAGACTGGATCTGTGTGACAAATGCGACCGTGA	1885	

## RESULT 9

US-10-007-926A-433  
; Sequence 433, Application US/10007926A  
; Publication No. US20030143539A1  
; GENERAL INFORMATION:  
; APPLICANT: BERTUCCI, FRANCOIS  
; APPLICANT: HOULGATTE, REMI  
; APPLICANT: BIRNBAUM, DANIEL

```

1  APPLICANT: NGUYEN, CATHERINE
2  APPLICANT: VIENS, PATRICIE
3  APPLICANT: FERT, VINCENT
4  TITLE OF INVENTION: GENE EXPRESSION PROFILING OF PRIMARY BREAST CARCINOMAS
5  TITLE OF INVENTION: USING ARRAYS OF CANDIDATE GENES
6  FILE REFERENCE: 1546-R-00
7  CURRENT APPLICATION NUMBER: US/10/007,926A
8  CURRENT FILING DATE: 2001-12-07
9  PRIOR APPLICATION NUMBER: 60/254,090
10 PRIOR FILING DATE: 2000-12-08
11 NUMBER OF SEQ ID NOS: 468
12 SOFTWARE: PatentIn Ver. 2.1
13 SEQ ID NO 433
14 LENGTH: 329
15 TYPE: DNA
16 ORGANISM: Homo sapiens
17 FEATURE:
18 OTHER INFORMATION: 5' terminal sequence. Homo sapiens plasminogen activator (PLAT)
19 FEATURE:
20 NAME/KEY: modified_base
21 LOCATION: (57)
22 OTHER INFORMATION: a, t, c or g
23 FEATURE:
24 NAME/KEY: modified_base
25 LOCATION: (82)
26 OTHER INFORMATION: a, t, c or g
27 FEATURE:
28 NAME/KEY: modified_base
29 LOCATION: (159)
30 OTHER INFORMATION: a, t, c or g
31 FEATURE:
32 NAME/KEY: modified_base
33 LOCATION: (264)
34 OTHER INFORMATION: a, t, c or g
35 FEATURE:
36 NAME/KEY: modified_base
37 LOCATION: (278)
38 OTHER INFORMATION: a, t, c or g
39 US-10-007-926A-433

```

## RESULT 10

RESUBMIT TO  
 US-09-880-503-15  
 ; Sequence 15, Application US/09880503  
 ; Patent No. US20020131964A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: CINES, Douglas B  
 ; APPLICANT: HITGAZI, Abd Al-Roof  
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND  
 ; TITLE OF INVENTION: TISSUE CONTRACTABILITY  
 ; FILE REFERENCE: 9596-331  
 ; CURRENT APPLICATION NUMBER: US/09/880,503  
 ; CURRENT FILING DATE: 2001-06-13



;; PRIOR APPLICATION NUMBER: US 60/212,847  
;; PRIOR FILING DATE: 2000-06-20  
;; NUMBER OF SEQ ID NOS: 18  
;; SOFTWARE: PatentIn Ver. 2.1  
;; SEQ ID NO 15  
;; LENGTH: 1212  
;; TYPE: DNA  
;; ORGANISM: Homo sapiens  
US-09-880-503-15

Query Match 18.0%; Score 191.6; DB 10; Length 1212;  
Best Local Similarity 52.3%; Pred. No. 1.5e-50;  
Matches 549; Conservative 0; Mismatches 459; Indels 42; Gaps 4;

QY 18 CTGCTACTTTGGGAATGGTCCAGCTACCGTGGCAGCAGCAGCCTCACCGAGTCGGGTGC 77  
DB 147 CTGCTATAGGGGAATGTCTACTTTACCGAGGAAGGCCAGCTGACACCATGGGCGC 206  
QY 78 CTCTGCTCCCTCGGTGGAATTCATGATCCTGATAGGCAAGGTTTACACAGCACAGAACCC 137  
DB 207 GCCCTGCTGCCCTGGAACTCTGCCACTGTCTTTCAGCAACAGTACCATGCCACAGATC 266  
QY 138 CAGTGCCAGGACACTGGGCTGGGCAACATAATTAATCTGCGGGAATCTGATGGGGATGC 197  
DB 267 TGATGCTCTTCAGCTGGGCTGGGGAACATAATTAATCTGCGGGAACCCAGACAAACCGGAG 326  
QY 198 CAAGCCCTGGTGCACGCTGGAACCGCAGGCTGACGTGGGAGTACTGTGATGTGC 257  
DB 327 GGGACCTTGTGCTATGTGAGGTGGGCTTAAGCCGCTTGTCCAGAGTGCATGGTGCA 386  
QY 258 CTCTCTCCACCTGCGCCTGAGACAGTACAGCCA-----GCCTCAGTT 302  
DB 387 TGACTGCGCAGATGAAAAATTAATAATTTTCAGTGTGGCCAAAAGACTCTGTGAGGCCCGCTT 446  
QY 303 TCGCATCAAGAGAGGGCTCTTCGCGGACATGCGCTCCACCCCTGGCAGGCTGCATCTT 362  
DB 447 TAAGATTATTGGGGAGAAATTCACCAACATCGAAGCAACGAGCCCTGGTTTGGCGGCATCTA 506  
QY 363 TGCCAAAGACAGGAGGTGCCCGGAGAGCGGTCTCTGTCGGGGGCATCTCATCAGTGC 422  
DB 507 CAGGAGGACCGGGGGGCTCTGTCA----CTACGTGTGTGAGGAGCAGCTCATCAGCCC 563  
QY 423 CTGCTGGAATCTCTGCGGCCACTGCTTCCAGGAGAGGTTTCGCGCCCAACCACTGAC 482  
DB 564 TTGCTGGGTGATCAGCGCCACACACTGCTTCAATTGATTACCAAGAGGAGGACTACAT 623  
QY 483 GGTGATCTTGGCGAGAACATACCGGTGCTCCTGGCGAGGAGGAGAGAAATTTGAGT 542  
DB 624 CGTCTACCTGGTCTCAAGGCTTAACTCCAACACGCAAGGGAGATGAAGTTTGAAGT 683  
QY 543 CGAAAAATACATTTGCCATAAGGAATTCGATGATGACA-----CTTACGACAATGACAT 596  
DB 684 GGAATACTCATCTACACAGGACTACAGCGCTGACACGCTTGTCTACCAACAGGACAT 743  
QY 597 TCGCTGTGAGCTGAAATCGGAATTCGTCCGCTGTGTCGCCAGGAGAGCAGCGTGGTCCG 656  
DB 744 TGCTTGTGATGATCCGTTCCAGGAGGCGAGGTGTGCGCAGCATCCCGGACTATACA 803  
QY 657 CACTGTGTGCTTCCCGCGGAGACTGCGAGTGCAGACTGCGGAGTGTGAGTGTGAGTCTC 716  
DB 804 GACCATCTGCTGCTCGATGTATACGATCCCGAGTTTGGCAACAAGCTGTGAGATCAC 863  
QY 717 CGGCTACCGCAGCATGAGGCTTGTCTCTCTTCTATTTCGAGCGGCTGAGGAGGCTCA 776  
DB 864 TGGCTTTGGAAGAAGAAATTTACCGACTATCTTATCCGAGCAGCTGAAAAATGACTGT 923  
QY 777 TGTCAAGCTGTACCCATCCAGCGCTGCACATCAACAATTTACTTTAACAGAACAGTCA 836  
DB 924 TGTGAAGCTGATTTCCACCGGAGTGTACAGCAGCCCACTACTACGCTCTGGAAGTCA 983  
QY 837 CGACAAATGTGTGTGTGAGACACTCGGAGCGGGCGGCCCAAGCAAACTTGCACGA 896  
DB 984 CACCAAAATGTGTGTGTGTGACACCC-----AATGGAACACAGA 1025

QY 897 CGCTGCGCAGGGGATTCGGGAGGCCCCCTGTGTGTGTGAACGATGCGCGCATGACTTT 956  
DB 1026 TTCTTCCAGGAGACTCAGGGGAGCCCCCTCGTGTGTTCCTCCAGGCCGATGACTTT 1085  
QY 957 GGTGGGATCATCAGCTGGGGCTGGGCTGTGACAGAGATGTCGGGGTGTGTACAC 1016  
DB 1086 GACTGGAATTTGTGAGCTGGGGCTGGATGTGCCCTGAAGCAACAGCGGCTGTACAC 1145  
QY 1017 AAAGTTTACCACTTACCTAGACTGGATTGC 1046  
DB 1146 GAGAGTCTCACACTTCTTACCCTGGATCCG 1175

RESULT 11

US-10-198-846-12748  
; Sequence 12748, Application US/10198846  
; Publication No. US2003009974A1  
; GENERAL INFORMATION:  
; APPLICANT: Lillie, James  
; APPLICANT: Xu, Yongyao  
; APPLICANT: Wang, Youzhen  
; APPLICANT: Steimann, Kathleen  
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS  
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND  
; FILE REFERENCE: W01-049  
; CURRENT APPLICATION NUMBER: US/10198,846  
; PRIOR FILING DATE: 2002-07-18  
; PRIOR APPLICATION NUMBER: 60/306,220  
; PRIOR FILING DATE: 2001-07-18  
; NUMBER OF SEQ ID NOS: 14084  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 12748  
; LENGTH: 1415  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: 1, 2, 3, 1414, 1415  
; OTHER INFORMATION: n = A,T,C or G  
US-10-198-846-12748

Query Match 17.4%; Score 185; DB 14; Length 1415;  
Best Local Similarity 92.6%; Pred. No. 2e-48;  
Matches 249; Conservative 0; Mismatches 15; Indels 5; Gaps 5;

QY 1 TCTGAGGAAACAGTGAAGTCTTCTTGGGAATGGGTGAGCTACCGTGGCACGACAGC 60  
DB 906 TCTGAGGAAACAGTGAAGTCTTCTTGGGAATGGGTGAGCTACCGTGGCACGACAGC 963  
QY 61 CTACCCAGTGGGTGCTCCT-GCCTCCGCTGGAATTCATGATCCTGATAGGCAAGT 119  
DB 964 CTACCCAGTGGGTGCTCCTCGCTCCGCTGGAATTCATGATCCTGATAGGCAAGT 1023  
QY 120 TTACACAGCAGACAGAACCCAGTGCAGGACCTGGGCTGGGCAACATAATTAATGCTCCG 179  
DB 1024 TTACACAGCAGACAGAACCCAGTGCAGGACCTGGGCTGGGCAACATAATTAATGCTCCG 1083  
QY 180 GAATCTGTGAGGGAATGCCAAGCCCTGGTGCACGCTGCTGAA-GAACCGCAGGCTGACGT 238  
DB 1084 GAAT-CTGATGGGATGCCAAGCCCTGTTTGCAGTCTGTGAAACGAAACCGCAGGCTGACGT 1142  
QY 239 GGGAGTACTGTGATGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 267  
DB 1143 GGGAGTACTGTGATGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1171

RESULT 12

US-09-735-705-122  
; Sequence 122, Application US/09735705  
; Patent No. US20020052329A1  
; GENERAL INFORMATION:







Db 1181 -----GAAAAAGATTTCTCCAGGAGACTCAGGGGACCCCTCGTCTG 1225  
Qy 933 TCTGAACGATGCGCGCATGCTTTGGTGGGCATCATCAGCTGGGGCTTGGGCTGTGGACA 992  
Db 1226 TTCCCTCCAGCGCGCATGCTTTGACTGGAAATGTGAGCTGGGCGCTGGATGTGCCCT 1285  
Qy 993 GAAGGATGTCGGGGTGTGACACAAAGGTTACCAACTACCTAGACTGGATTG 1046  
Db 1286 GAAGGACAAAGCCAGGCGTCTACAGGAGTCTCACACTTCTTACCTGGATCCG 1339

## RESULT 15

US-09-466-396A-122

; Sequence 122, Application US/09466396A

; Publication No. US20030119763A1

; GENERAL INFORMATION:

; APPLICANT: Wang, Tongtong

; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY AND

; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER

; FILE REFERENCE: 210121.455C4

; CURRENT APPLICATION NUMBER: US/09/466.396A

; CURRENT FILING DATE: 1999-12-17

; NUMBER OF SEQ ID NOS: 224

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 122

; LENGTH: 1475

; TYPE: DNA

; ORGANISM: Homo sapien

US-09-466-396A-122

Query Match 16.0%; Score 170.8; DB 11; Length 1475;

Best Local Similarity 51.3%; Pred. No. 7e-44;

Matches 551; Conservative 0; Mismatches 457; Indels 66; Gaps 4;

Qy 18 CTGCTACTTTGGGAATGGTCCAGCTACCGTGGGACGACAGCTCACCAGTCCGGTGC 77  
Db 287 CTGCTATGAGGGGAATGGTCACTTTTACCGAGGAAAGCCAGCACTGACCATGGGCGG 346  
Qy 78 CTCCTGCTCCCGTGGGAATTCATGATCCTGATAGGCAAGGTTTACACAGCACAGAACC 137  
Db 347 GCCCTGCTGCCCTGGAACTCTGCCACTGTCTTCCAGAAAGTACCATGCCACAGATC 406  
Qy 138 CAGTGCCAGGCACTGGGCTGGGCAACAATAATTACTGCGGAATCCTGATGGGGATGC 197  
Db 407 TGATGCTCTTCAGCTGGGCTGGGCAACAATAATTACTGCGGAATCCTGATGGGGATGC 466  
Qy 198 CAAGCCCTGGTGGCCAGCTGTGAAGAACCGGAGCTGACGTGGGAGTACTGTGATGCC 257  
Db 467 GCGACCTGGTGTATGTGACAGTGGGCTTAAAGCCGCTTGTCCAAGAGTGATGGTGCA 526  
Qy 258 CTCCTGCTCCACCTCGCGCCTGAG----- 281  
Db 527 TGACTGGCGAGATGGAAGAGCCCTCTCTCTCCAGAGAATAAATTCAGTGTGG 586  
Qy 282 ---ACAGTACAGCAGCCTCAGTTTCGCATCAAGAGGGGCTTTTCGCCGACATCGCCTC 338  
Db 587 CCAAAAGACTCTGAGGCCCGCTTTAAGATTATTGGGGAGAAATTCACCACCATCGAGAA 646  
Qy 339 CCACCCCTGGCAGGCTGCTATCTTTGCAAGCAGAGAGTTCGCCCGGAGAGCGGTTCT 398  
Db 647 CCAGCCCTGGTTTGGGCGCATCTACAGGAGGACCGGGG---GGGCTCTGTCACTACGT 703  
Qy 399 GTGCGGGGGATCTCATCAGCTCTGCTGGATTTCTCTGCGCGCCCACTGCTTCCAGGA 458  
Db 704 GTGTGGAGGAGCTCATCAGCCCTTGTCTGGGTGATCAGCCCAACACACTGCTTATTGA 763  
Qy 459 GAGGTTTCCGCCCAACCACTGACGGTGTCTTGGGAGAGAACATACCGGGTGTCCCTGG 518  
Db 764 TTACCCAAAGAGGAGGACTACATCGTCTACCTGGTGGCTCAAGGCTTAACCTCAACAC 823  
Qy 519 CGAGGAGGAGAGAAATTTGAAGTCGAAATAATTTGTCATTAAGGAATTCGATGATGA 578

Db 824 GCAAGGGGAGATGAAGTTTGAGGTGGAAAACTCATCTACACAAGGACTACAGCGCTGA 883  
Qy 579 CA-----CTTAGCAAAATGACATTTGCGTGTGCTGCAAGTGAATTCGATTCGCTCCGCTG 632  
Db 884 CACGCTTGCTCACCACAAACGACATTTGCTTGTGTAAGATCCGTTTCCAAGGAGGCGAGTG 943  
Qy 633 TGGCCAGGAGAGCAGCGTGTGCGCACTGTGTGCTTCCCTCCCGGCGGACCTGCACTGCC 692  
Db 944 TGGCGAGCCATCCCGGACTATACAGACCATCTGCTGCTGCTCGATGTATTAACGATCCCA 1003  
Qy 693 GGAATGAGGAGTGTGAGCTCTCCGCTACGCAAGCATGAGGCGCTTGTCTCTTTCTA 752  
Db 1004 GTTTGGCAAGCTGTGAGATCACTGGCTTTGAAAAGAAATTTTACCGACTATCTTA 1063  
Qy 753 TTGGAGCGGCTGAAGGAGGCTCATGTGCACTGTATCCCATCCAGCGCTGCAATCACA 812  
Db 1064 TCCGGAGCAGCTGAAGATGACTGTGTGAAGCTGATTTCCACCGGAGTGTCAAGCAGCC 1123  
Qy 813 ACATTTACTTAAAGAACAGTCAACGACATGCTGTGTGTGTGGAGACACTCGGAGCGG 872  
Db 1124 CCACCTACTACGGCTCTGAAGTCAACCAAAATGCTGTGTGTGTGCTGACCCACAGTG- 1180  
Qy 873 CGGGCCCCAGGCAAACTTGCACGACGCTGCGAGGCGATTGCGGAGGCGCCCTGTGTG 932  
Db 1181 -----GAAAAAGATTTCTCCAGGAGACTCAGGGGACCCCTCGTCTG 1225  
Qy 933 TCTGAACGATGCGCGCATGCTTTGTGGGCACTCATCAGCTGGGGCTGCGGTGTGGACA 992  
Db 1226 TTCCCTCCAAGGCGCGCATGACTTTGACTGGAATTTGTCAGCTGGGGCGGTGATGTGCCCT 1285  
Qy 993 GAAGGATGTCGGGGTGTGTACACAAGGTTTACAACTACTAGACTGGATTG 1046  
Db 1286 GAAGGACAAAGCCAGGCGTCTACACGAGAGTCTCACACTTCTTACCCCTGGATCCG 1339

Search completed: August 19, 2003, 14:22:30

Job time : 302.529 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 18, 2003, 22:44:33 ; Search time 3115.64 Seconds  
(without alignments)  
8307.845 Million cell updates/sec

Title: US-09-987-455-7

Perfect score: 1065

Sequence: 1 tctgagggaacagtgactg.....gtgacaacatcgaccgtga 1065

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST.\*

1: em\_estba.\*

2: em\_esthum.\*

3: em\_estin.\*

4: em\_estmu.\*

5: em\_estov.\*

6: em\_estpl.\*

7: em\_estro.\*

8: em\_estl.\*

9: gb\_estl.\*

10: gb\_est2.\*

11: gb\_estc.\*

12: gb\_est3.\*

13: gb\_est4.\*

14: gb\_est5.\*

15: em\_estfun.\*

16: em\_estom.\*

17: em\_gss\_hum.\*

18: em\_gss\_inv.\*

19: em\_gss\_pln.\*

20: em\_gss\_vrt.\*

21: em\_gss\_fun.\*

22: em\_gss\_mam.\*

23: em\_gss\_mus.\*

24: em\_gss\_pro.\*

25: em\_gss\_rod.\*

26: em\_gss\_phg.\*

27: em\_gss\_vrl.\*

28: gb\_gsal.\*

29: gb\_gsal.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	895	84.0	930	13	BX329047
2	853	80.1	912	13	BUI49958
3	850.6	79.9	962	13	BX389609
4	818.6	76.9	891	13	BUI46167

5	815.8	76.6	993	13	BQ278177
6	811.2	76.2	883	13	BQ690978
7	807.6	75.8	918	13	BX389608
8	796.2	74.8	930	13	BUI45014
9	782.8	73.5	898	13	BUI57720
10	767.4	72.1	916	12	B1765247
11	765.2	71.8	950	13	BQ927899
12	706.8	66.4	948	10	BEG16613
13	697.4	65.5	958	12	BG763582
14	694.4	65.2	974	13	BQ960216
15	693.4	65.1	709	10	BE732704
16	685.2	64.3	1060	12	BM556999
17	677	63.6	1042	10	BE732413
18	672	63.1	957	13	BQ687538
19	668.8	62.8	948	13	BQ687779
20	665.2	62.5	704	9	AUI134301
21	658.2	61.8	903	13	BQ689095
22	656.2	61.6	830	13	BQ896938
23	655.2	61.5	904	13	BQ689840
24	655.2	61.5	911	13	BQ684734
25	654.6	61.5	867	13	BUI79903
26	652.4	61.3	818	9	AUI124602
27	651	61.1	818	12	BM006474
28	649.6	61.0	757	13	BQ879911
29	642.2	60.3	1037	13	BUI57493
30	631.2	59.3	813	10	BG120780
31	629.8	59.1	663	9	AUI134463
32	624	58.6	624	10	BE549489
33	595.4	55.0	587	10	BE394944
34	561	52.7	561	9	AL047284
35	553.4	52.0	555	14	CB128855
36	535.4	50.3	537	14	CB155052
37	531	49.9	964	12	BG769384
38	517.4	48.6	905	13	BUI59271
39	514.8	48.3	893	14	CD359959
40	514.4	48.3	516	14	CB140923
41	496.4	46.6	916	13	BUI57241
42	487.8	45.8	527	9	AW630119
43	485.8	45.6	553	9	AW394127
44	477.8	44.9	917	13	BUI190904
45	475.8	44.7	937	13	BQ679501

#### ALIGNMENTS

RESULT 1  
BX329047/c  
LOCUS  
DEFINITION  
BX329047 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens CDNA  
clone CSODI039YK15 3-PRIME, mRNA sequence.

ACCESSION  
BX329047

VERSION  
BX329047.1

KEYWORDS  
EST.

SOURCE  
Homo sapiens (human)

ORGANISM  
Homo sapiens

REFERENCE  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

TITLE  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS  
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.

COMMENT  
Full-length cDNA libraries and normalization

Unpublished

Contact: Genoscope

Genoscope - Centre National de Sequencage

BP 191 91006 EVRY cedex - France

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr

Library was constructed by Life Technologies, a division of

Invitrogen. This sequence belongs to sequence cluster 4498.r For

more information about this cluster, see

http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CSOBAI0212D08\_C

S01980\_tcluster=4498.r. Contact : Feng Liang Email :

fliang@lifetech.com URL : http://fulllength.invitrogen.com/

Invitrogen Corporation 1600 Faraday Avenue Genoscope sequence ID :

```

CSOBAI021ZD08 CS01980 1.
FEATURES
  source
    Location/Qualifiers
      1..930
        /organism="Homo sapiens"
        /mol_type="mRNA"
        /db_xref="taxon:9606"
        /clone="CSODI039YK15"
        /tissue_type="PLACENTA COT 25-NORMALIZED"
        /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
        /note="1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT
  184 a 268 c 282 g 194 t 2 others
ORIGIN
  Query Match 84.0%; Score 895; DB 13; Length 930;
  Best Local Similarity 98.6%; Pred. No. 5.7e-223;
  Matches 912; Conservative 0; Mismatches 12; Indels 1; Gaps 1;
QY 1 TCTGAGGGAACAGTACTTCTTGGGAATGGTCAAGCTACCGTGGCAGCACAGC 60
Db 927 TTTGAGGGAACAGTACTTCTTGGGAATGGTCAAGCTACCGTGGCAGCACAGC 869
QY 61 CTCACCGAGTCGGTGCCTCTCGCTCCCGTGGAAATCCATGATCCTGATAGGCAAGTT 120
Db 868 CTCACCGAGTCGGTGCCTCTCGCTCCCGTGGAAATCCATGATCCTGATAGGCAAGTT 809
QY 121 TACACAGACAGAACCCAGTCGCCAGCACTGGGCTGGGCAAAACATAATATTCTGCCGG 180
Db 808 TACACAGACAGAACCCAGTCGCCAGCACTGGGCTGGGCAAAACATAATATTCTGCCGG 749
QY 181 AATCTGATGGGATGCCAAGCCCTGTGTCACCTGCTGAGAACCCGAGCTCACTGG 240
Db 748 AATCTGATGGGATGCCAAGCCCTGTGTCACCTGCTGAGAACCCGAGCTCACTGG 689
QY 241 GAGTACTGTGATGTCCCTCTCTGCTCCACCTCGCGCTGACAGTAGTACAGCCCTCAG 300
Db 688 GAGTACTGTGATGTCCCTCTCTGCTCCACCTCGCGCTGACAGTAGTACAGCCCTCAG 629
QY 301 TTTCCGATCAAGAGAGGCTCTTCCCGACATCCCTCCACCCCTGGCAGGCTGCCATC 360
Db 628 TTTCCGATCAAGAGAGGCTCTTCCCGACATCCCTCCACCCCTGGCAGGCTGCCATC 569
QY 361 TTTCCGACACAGAGGTCCCGGAGAGGCTTCTGTGCGGGGACATCTCATCAGC 420
Db 568 TTTCCGACACAGAGGTCCCGGAGAGGCTTCTGTGCGGGGACATCTCATCAGC 509
QY 421 TCCTGCTGGATTCTCTCTGCGCGCCACTGCTTCAGAGAGAGTTTCGCGCCCAACCTG 480
Db 508 TCCTGCTGGATTCTCTCTGCGCGCCACTGCTTCAGAGAGAGTTTCGCGCCCAACCTG 449
QY 481 ACGTGATCTTTGGGAGAAACATACCGGCTGCTCCCTGGCGAGGAGGAGCAAAATTTGAA 540
Db 448 ACGTGATCTTTGGGAGAAACATACCGGCTGCTCCCTGGCGAGGAGGAGCAAAATTTGAA 389
QY 541 GTCCAAAATACATGTTCATAGGAATTCATATGACATTCAGCAATGCAATTCGCG 600
Db 388 GTCCAAAATACATGTTCATAGGAATTCATATGACATTCAGCAATGCAATTCGCG 329
QY 601 CTGTGACGCTGAAATCGATTCTGCTCCGCTGTGCCAGAGAGCAGCGTGGTCCGCACT 660
Db 328 CTGTGACGCTGAAATCGATTCTGCTCCGCTGTGCCAGAGAGCAGCGTGGTCCGCACT 269
QY 661 GTGTGCTTCCCGCGGAGCCTCGAGCTGCCGACTGGACGAGTGTGAGCTCTCCGGC 720
Db 268 GTGTGCTTCCCGCGGAGCCTCGAGCTGCCGACTGGACGAGTGTGAGCTCTCCGGC 209
QY 721 TACGCGACATGAGGCTTCTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 780
Db 208 TACGCGACATGAGGCTTCTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 149
QY 781 AGACTGTACCCATCCAGCGCTGCACATCAACAATTTACTTTAAAGAACAGTCAGTCAACGAC 840

```

---

```

148 AGACTGTACCCATCCAGCGCTGCACATCAACAATTTACTTTAAAGAACAGTCAGTCAACGAC 89
841 AACATGCTGTGTGCTGTGAGACACTCGAGCGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 900
88 AACATGCTGTGTGCTGTGAGACACTCGAGCGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 29
901 TGCACGGGGGATTCGGGAGGCGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 925
28 TGCACGGGGGATTCGGGAGGCGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 4

RESULT 2
LOCUS
DEFINITION
  BU149958 912 bp mRNA linear EST 03-SEP-2002
  AGENCOURT 8137213 Lupski_dorsal_root_ganglion Homo sapiens cDNA
  clone IMAGE:6184119 5', mRNA sequence.
ACCESSION
  BU149958
VERSION
  BU149958.1 GI:22663490
KEYWORDS
  EST.
SOURCE
  Homo sapiens (human)
  Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
  1 (bases 1 to 912)
  NIH-MGC http://mgi.nci.nih.gov/
  National Institutes of Health, Mammalian Gene Collection (MGC)
  Unpublished
  Contact: Robert Strausberg, Ph.D.
  Email: cgabs-r@mail.nih.gov
  Tissue Procurement: Dr. James R. Lupski
  cDNA Library Preparation: Life Technologies, Inc.
  cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
  DNA Sequencing by: Agencourt Bioscience Corporation
  Clone distribution: MGC clone distribution information can be
  found through the I.M.A.G.E. Consortium/LLNL at:
  http://image.llnl.gov
  Plate: LLAM13572 row: m column: 16
  High quality sequence start: 18
  High quality sequence stop: 644.
FEATURES
  source
    Location/Qualifiers
      1..912
        /organism="Homo sapiens"
        /mol_type="mRNA"
        /db_xref="taxon:9606"
        /clone="IMAGE:6184119"
        /sex="male"
        /tissue_type="dorsal root ganglia"
        /dev_stage="adult, 36 yr"
        /lab_host="DH10B"
        /clone_lib="Lupski dorsal root ganglion"
        /notes="Vector: pCMV-SPORT6 (Life Technologies); Site_1:
        NotI; Site_2: SalI; cDNA made by oligo-dT priming.
        Directionally cloned using the following adaptors:
        5'-TCGACCCAGCGCTCCG-3' and
        5'-GACTAGTTCCTAGATCGAGCGCGCCCT(15)-3'. Size selected >
        1 kb for average insert length 1.7 kb. This is a primary
        library, non-amplified. Library constructed by Life
        Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
        College of Medicine) and is available through Life
        Technologies."
BASE COUNT
  189 a 272 c 271 g 179 t 1 others
ORIGIN
  Query Match 80.1%; Score 853; DB 13; Length 912;
  Best Local Similarity 98.9%; Pred. No. 5.4e-212;
  Matches 901; Conservative 0; Mismatches 5; Indels 5; Gaps 4;
QY 131 AGAACCCCGTGGCGGACCTGGG--CCTGGCAACATAATTAATCTGCCGAATCTGA 188
Db 1 AGAACCCCGTGGCGGACCTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 60
QY 189 TGGGAGTGCACACCCCTGTGTGCGCAGTGTGTGAAGAACCGCAGGCTGACGTGGGAGTACTG 248

```

```

Db 61 TGGGGATGCCAAGCCCTGGTCCACGTCCTGAAGAACCGCAGGCTGACGTCGGAGTACTG 120
Qy 249 TGATGTGCGCCCTCCCTGCTCCACCTGCGGCTCAGACAGTACAGCCAGCCTCAGTTTCGCAT 308
Db 121 TGATGTGCGCCCTCCCTGCTCCACCTGCGGCTCAGACAGTACAGCCAGCCTCAGTTTCGCAT 180
Qy 309 CAAAGGAGGCTCTTTCGCGCATCGCTCCACCTCCGAGGCTGCGGCTGCGGCTGCGGCTG 368
Db 181 CAAAGGAGGCTCTTTCGCGCATCGCTCCACCTCCGAGGCTGCGGCTGCGGCTGCGGCTG 240
Qy 369 GCACAGAGGTCGCGCGGAGGCGGTTCTCTGTCGGGGGATATCATCAGCTCTGCTG 428
Db 241 GCACAGAGGTCGCGCGGAGGCGGTTCTCTGTCGGGGGATATCATCAGCTCTGCTG 300
Qy 429 GATTCTCTCTGCGCGGCTCTCTTCCAGGAGGTTTCCGCGCCACACCTGACGCGTAT 488
Db 301 GATTCTCTCTGCGCGGCTCTCTTCCAGGAGGTTTCCGCGCCACACCTGACGCGTAT 360
Qy 489 CTTGGGAGAAATACATACCGGGTGGTCCCTGGCGAGGAGAGCAGAAATTTGAAGTCGAAA 548
Db 361 CTTGGGAGAAATACATACCGGGTGGTCCCTGGCGAGGAGGAGCAGAAATTTGAAGTCGAAA 420
Qy 549 ATACATTGTCATTAAGNAATTCGATGATGACATTCATCAGCAATGACATGCGCTGCTGCA 608
Db 421 ATACATTGTCATTAAGNAATTCGATGATGACATTCATCAGCAATGACATGCGCTGCTGCA 480
Qy 609 GCTGAATTCGATTCGCTCCGCTGTCGCGAGGAGGCTGTCGCGACGCTGTCGCTGCTGCT 668
Db 481 GCTGAATTCGATTCGCTCCGCTGTCGCGAGGAGGCTGTCGCGACGCTGTCGCTGCTGCT 540
Qy 728 TCCCGCGCGGACCTGACGCTGCGGAGTGGAGCGAGTGTGAGCTCTCCGCTACGGCAA 728
Db 541 TCCCGCGCGGACCTGACGCTGCGGAGTGGAGCGAGTGTGAGCTCTCCGCTACGGCAA 600
Qy 729 GCATGAGCCTTGTCTCTTCTTATTTGGAGCGGCTGAGAGGCTCATGTCAGACTGTA 788
Db 601 GCATGAGCCTTGTCTCTTCTTATTTGGAGCGGCTGAGAGGCTCATGTCAGACTGTA 660
Qy 789 CCATTCAGCGCTGCAATCACAACATTTACTTTAAACAGACAGTCAACGACACATGCT 848
Db 661 CCATTCAGCGCTGCAATCACAACATTTACTTTAAACAGACAGTCAACGACACATGCT 720
Qy 849 GTGTGCTGGAGACTCGAGCGCGGCGCCAGGCAAACTTGCACGACGCTCGCAGGG 908
Db 721 GTGTGCTGGAGACTCGAGCGCGGCGCCAGGCAAACTTGCACGACGCTCGCAGGG 780
Qy 909 -CGATTTCGGAGG-CCCGCTGCTGTCTGAACGATGCGCGCATGCTTTGGTGGGATC 966
Db 781 NCGATTTCGGAGGCGCGCCCTGCTGTCTGAACGATGCGCGCATGCTTTGGTGGGATC 840
Qy 967 ATCAGCTGGGCGCTGGGCTGGGACAGAGGATGT-CCCGGCTGTGTACACAAAGGTTAC 1025
Db 841 ATCAGCTGGGCGCTGGGCTGGGACAGAGGATGTCCCCGGGTGTGTACACAAAGGTTAC 900
Qy 1026 CAATACCTAG 1036
Db 901 AAATACCTAG 911

```

## RESULT 3

BX389609

LOCUS

DEFINITION BX389609 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA

Clone CS0D1039YK15 5-PRIME, mRNA sequence.

ACCESSION BX389609

VERSION BX389609.1 GI:30463470

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 962)

## AUTHORS

Li, W.B., Gruber, C., Jessee, J. and Polayes, D.

Full-length cDNA libraries and normalization

## JOURNAL

Unpublished

## COMMENT

Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 EVRY cedex - France  
Email: seqr@genoscope.cns.fr, Web : www.genoscope.cns.fr  
was normalized. Library was constructed by Life Technologies, a  
division of Invitrogen. This sequence belongs to sequence cluster  
4498.r For more information about this cluster, see  
http://www.genoscope.cns.fr/

cgi-bin/cluster.cgi?seq=CS0BAF025ZG09\_AF02376\_3&cluster=4498.r.  
Contact : Feng Liang Email : fliang@lifetech.com URL :  
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600  
Paradise Avenue Genoscope sequence ID : CS0BAF025ZG09\_AF02376\_3.

## FEATURES

Location/Qualifiers

source.

1..962

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="CS0D1039YK15"

/tissue\_type="PLACENTA COT 25-NORMALIZED"

/clone\_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"

/notes="1st strand cDNA was primed with a NotI-oligo(dT)

primer. Five prime end enriched, double-strand cDNA was

digested with Not I and cloned into the Not I and EcoR V

sites of the pCMVSPORT 6 vector. Library was normalized."

BASE COUNT 203 a 290 c 277 g 191 t 1 others

## ORIGIN

Query Match 79.9%; Score 850.6; DB 13; Length 962;

Best Local Similarity 96.4%; Pred. No. 2.3e-211;

Matches 891; Conservative 0; Mismatches 30; Indels 3; Gaps 2;

Qy 142 GCCCAGGCACTGGCCCTGGGCAACATAATTAATCTGCCGGAATCTGTATGGGATGCCAAG 201

Db 1 GCCCAGGCACTGGGCGCTGGGCAAACTAAATTAATCTGCCGGAATCTGTATGGGATGCCAAG 60

Qy 202 CCTGTGTGCCACGTGTCTGAAGAACCGCAGGCTGACGCTGGGAGTACTGTGTATGTCCTCC 261

Db 61 CCTGTGTGCCACGTGTCTGAAGAACCGCAGGCTGACGCTGGGAGTACTGTGTATGTCCTCC 120

Qy 262 TGCTTCCACCTGCGGCTTGAGACAGTACAGCAGCCTCAGTTTCGATCAAAGGAGGCTC 321

Db 121 TGCTTCCACCTGCGGCTTGAGACAGTACAGCAGCCTCAGTTTCGATCAAAGGAGGCTC 180

Qy 322 TTGCCCGACATCGCCTCCACCCCTGGCAGGCTGCCATCTTTGCCACGACACAGAGGCTC 381

Db 181 TTGCCCGACATCGCCTCCACCCCTGGCAGGCTGCCATCTTTGCCACGACACAGAGGCTC 240

Qy 382 CCGGAGAGCGGTTCTGTGCGGGGGCATACTCATCAGCTCTCTGCTGATTTCTCTGCC 441

Db 241 CCGGAGAGCGGTTCTGTGCGGGGGCATACTCATCAGCTCTCTGCTGATTTCTCTGCC 300

Qy 442 GCCCAGCTTCCAGGAGAGGTTTCCGCCCCACCTGACGCTGATCTTTGGGAGAGCA 501

Db 301 GCCCAGCTTCCAGGAGAGGTTTCCGCCCCACCTGACGCTGATCTTTGGGAGAGCA 360

Qy 502 TACCGGTGTCTCTGGCGAGGAGGAGCAGAAATTTGAAGTCGAAAAATACATTTGTCAT 561

Db 361 TACCGGTGTCTCTGGCGAGGAGGAGCAGAAATTTGAAGTCGAAAAATACATTTGTCAT 420

Qy 562 AAGGAATTCATGATGACACTTACGACAATGACATTCGCTGCTGTCAGCTGAAATTCGAT 621

Db 421 AAGGAATTCATGATGACACTTACGACAATGACATTCGCTGCTGTCAGCTGAAATTCGAT 480

Qy 622 TCCTCCCGCTGTGCCAGGAGAGCAGCGTGGTCCGCACTGTGTGCTCTCCCGCGCGGAC 681

Db 481 TCCTCCCGCTGTGCCAGGAGAGCAGCGTGGTCCGCACTGTGTGCTCTCCCGCGCGGAC 540

Qy 682 CTGACGCTGCCGACTGACGAGTGTGAGCTCTCCGGCTACGGCAAGCATGAGGCTTG 741

Db 541 CTGACGCTGCCGACTGACGAGTGTGAGCTCTCCGGCTACGGCAAGCATGAGGCTTG 600

742 TCTCTTTCTATTTCGAGCGGCTGAAGAGGCTCATGTGAGACTGTACCATCCAGCGC 801  
 Db |||||  
 601 TCTCTTTCTATTTCGAGCGGCTGAAGAGGCTCATGTGAGACTGTACCATCCAGCGC 660  
 Qy |||||  
 802 TGCAATCACAACATTTACTTTAAACAGAACAGTCAACGACCAATGCTGTGTGGAGAC 861  
 Db |||||  
 661 TGCAATCACAACATTTACTTTAAACAGAACAGTCAACGACCAATGCTGTGTGGAGAC 720  
 Qy |||||  
 862 ACTGGAGCGCGCGCGCCAGGCAACTTGACAGCGCTGCCAGGGGATTCGGAGGC 921  
 Db |||||  
 721 ACTGGAGCGCGCGCGCCAGGCAACTTGACAGCGCTGCCAGGGGATTCGGAGGC 780  
 Qy |||||  
 922 CCCCTGGTGTGTGAACAGTGGCGCATGCTTTGGTGGGCATCATCAGCTGGGGCTG 981  
 Db |||||  
 781 CCCCTGGTGTGTGAACAGTGGCGCATGCTTTGGTGGGCATCATCAGCTGGGGCTG 840  
 Qy |||||  
 982 GGCTGTGACAGAGGATGTCGGGTGTGTGACAAAGGTTACCAACTAGACTGG 1041  
 Db |||||  
 841 GGCTGTGACAGAGGATGTCGGGTGTGTGACAAAGGTTACCAACTAGACTGG 898  
 Qy |||||  
 1042 ATTCTGACACATGCGACCTGA 1065  
 Db |||||  
 899 ATTCGGGAC-ACATGCGACCGGA 921

RESULT 4  
 BUI46167  
 LOCUS  
 DEFINITION AGENCOURT 7981262 lupski\_dorsal root ganglion Homo sapiens cDNA  
 clone IMAGE:6185661 5', mRNA sequence.  
 ACCESSION BUI46167  
 VERSION BUI46167.1 GI:22659699  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 891)  
 NIH-MGC <http://mgc.nci.nih.gov/>.  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished  
 Contact: Robert Strausberg, Ph.D.  
 Email: [cgapbs@mail.nih.gov](mailto:cgapbs@mail.nih.gov)  
 Tissue Procurement: Dr. James R. Lupski  
 cDNA Library Preparation: Life Technologies, Inc.  
 DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)  
 Clone Distribution: Agencourt Bioscience Corporation  
 found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
 Plate: LLAM13576 row: m column: 22  
 High quality sequence stop: 665.  
 Location/Qualifiers  
 1..891  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:6185661"  
 /sex="male"  
 /tissue\_type="dorsal root ganglia"  
 /dev\_stage="adult, 36 yr"  
 /lab\_hosts="DH10B"  
 /clone\_lib="Lupski dorsal root ganglion"  
 /note="Vector: pCMV-SPORT6 (Life Technologies); Site 1:  
 NotI; Site 2: SalI; cDNA made by oligo-dr priming.  
 Directionally cloned using the following adaptors:  
 5'-TCGACCCACGGGTCCG-3' and  
 5'-GACTAGTCTAGTCGGCGGCGCCT(15)-3'. Size selected >  
 1 kb for average insert length 1.7 kb. This is a primary  
 library, non-amplified. Library constructed by Life  
 Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor  
 College of Medicine) and is available through Life

## FEATURES

source  
 1..891  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:6185661"  
 /sex="male"  
 /tissue\_type="dorsal root ganglia"  
 /dev\_stage="adult, 36 yr"  
 /lab\_hosts="DH10B"  
 /clone\_lib="Lupski dorsal root ganglion"  
 /note="Vector: pCMV-SPORT6 (Life Technologies); Site 1:  
 NotI; Site 2: SalI; cDNA made by oligo-dr priming.  
 Directionally cloned using the following adaptors:  
 5'-TCGACCCACGGGTCCG-3' and  
 5'-GACTAGTCTAGTCGGCGGCGCCT(15)-3'. Size selected >  
 1 kb for average insert length 1.7 kb. This is a primary  
 library, non-amplified. Library constructed by Life  
 Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor  
 College of Medicine) and is available through Life

## RESULT 5

BQ278177  
 LOCUS  
 DEFINITION AGENCOURT\_7061709 NIH\_MGC\_109 Homo sapiens cDNA clone IMAGE:5805153

## Technologies."

BASE COUNT 178 a 266 c 271 g 176 t  
 ORIGIN  
 Query Match 76.9%; Score 818.6; DB 13; Length 891;  
 Best Local Similarity 98.2%; Pred. No. 5.2e-203;  
 Matches 849; Conservative 0; Mismatches 14; Indels 2; Gaps 2;  
 Qy 136 CCAGTGGCCAGGACATGGGCTGGCCCAACATAATTAATGTCGCGGAATCTGTATGGGGAT 195  
 Db 1 CCAGTGGCCAGGACATGGGCTGGCCCAACATAATTAATGTCGCGGAATCTGTATGGGGAT 60  
 Qy 196 GCCAAGCCCTGGTGCCACGCTGCAAGAACCGCAGGCTGACGTGGGAGTACTGTGATGTG 255  
 Db 61 GCCAAGCCCTGGTGCCACGCTGCAAGAACCGCAGGCTGACGTGGGAGTACTGTGATGTG 120  
 Qy 256 CCCTCTGTCTCACTGGGCTTGACAGACGTAAGCAGCCTCAGTTTCCGATCAAAAGGA 315  
 Db 121 CCCTCTGTCTCACTGGGCTTGACAGACGTAAGCAGCCTCAGTTTCCGATCAAAAGGA 180  
 Qy 316 GGGCTCTTGGCCGACATGGCTCCCAACCTGGCAGGCTGCCATCTTTGCCAAGCAGG 375  
 Db 181 GGGCTCTTGGCCGACATGGCTCCCAACCTGGCAGGCTGCCATCTTTGCCAAGCAGG 240  
 Qy 376 AGGTGCGCCGAGAGCGGTTCTGTGCGGGGGCATACTCATCAGCTCCTGTGATTTCTC 435  
 Db 241 AGGTGCGCCGAGAGCGGTTCTGTGCGGGGGCATACTCATCAGCTCCTGTGATTTCTC 300  
 Qy 436 TGTGCGCCCACTGTTTCCAGGAGAGGTTTCCGCCCAACCACTGACGGTGTATTTGGGC 495  
 Db 301 TGTGCGCCCACTGTTTCCAGGAGAGGTTTCCGCCCAACCACTGACGGTGTATTTGGGC 360  
 Qy 496 AGAACATACCGGGTGGTCCCTGGCGGAGGAGCAGAAATTTGAAGTCGAAAATACATT 555  
 Db 361 AGAACATACCGGGTGGTCCCTGGCGGAGGAGCAGAAATTTGAAGTCGAAAATACATT 420  
 Qy 556 GTCCATAAGGAATTCGATGATGACATTTAGACAATGACATTTGGCTGTGCTGCACTGAAA 615  
 Db 421 CTCCATTAAGGAATTCGATGATGACATTTAGACAATGACATTTGGCTGTGCTGCACTGAAA 480  
 Qy 616 TCGGATTTGTCCTGCTGTGTCAGGAGAGCAGCGTGTCCGCACTGTGTGCTTCCCGCG 675  
 Db 481 TCGGATTTGTCCTGCTGTGTCAGGAGAGCAGCGTGTCCGCACTGTGTGCTTCCCGCG 540  
 Qy 676 GCGGACTGCGAGCTGCGGACTGGAAGTGTGAGCTTCCGCTACGCGCAGCATGAG 735  
 Db 541 GCGGACTGCGAGCTGCGGACTGGAAGTGTGAGCTTCCGCTACGCGCAGCATGAG 600  
 Qy 736 GCCTTGTCTCTTTCTATTTCGGAGCGGCTGAAGGAGGCTCATGTGACAGTGTACCCATCC 795  
 Db 601 GCCTTGTCTCTTTCTATTTCGGAGCGGCTGAAGGAGGCTCATGTGACAGTGTACCCATCC 660  
 Qy 796 AGCGGTGACATCACAACATTTACTTTAAACAGAACAGTCAACGACCAACATGCTGTGTGCT 855  
 Db 661 AGCGGTGACATCACAACATTTACTTTAAACAGAACAGTCAACGACCAACATGCTGTGTGCT 720  
 Qy 856 GGAGACACTCGGAGCGGC-GGGCCCCAGGCAACTTGCAGCAGCGCTGCCAGGGCGGATTC 914  
 Db 721 GGAGACACTCGGAGCGCGGGGGCCCCAGGCAAACTTGCAGCAGCGCTGCCAGGGCGGATTC 780  
 Qy 915 GGGAGGCCCCCTGTGTGTCTGAACGATGCCCGCAT-GACTTTGGTGGGCATCATCAGCT 973  
 Db 781 GGGAGGCCCCCTGTGTGTCTGAACGATGCCCGCATGAGACTTTGGTGGGCATCATCAGCT 840  
 Qy 974 GGGGCTTGGGCTGTGGACAGAGGA 998  
 Db 841 GGGGCGCCCTGGGCTGGGGGACAGGA 865







```

Db      497 TGGTCCGACATGTGTGCTTCCCCCGGCGACCTGACGTCCGACTGGACGGAGTGTG 556
Qy      710 AGCTCTCCGGCTACGCAAGCATGAGGCTTGTCTCTCTCTTCTATTCGAGCGGTGAAGG 769
Db      557 AGCTCTCCGGCTACGCAAGCATGAGGCTTGTCTCTCTCTTCTATTCGAGCGGTGAAGG 616
Qy      770 AGGCTCATGTGACATGTACCATCCAGCGCTGACATCAACATTTACTTAACAGAA 829
Db      617 AGGCTCATGTGACATGTACCATCCAGCGCTGACATCAACATTTACTTAACANGA 676
Qy      830 CAGTCACCGCAACAATGTGTGCTGGAGACACTCGAGCGGGGCCCGAGCAAACT 889
Db      677 CAGTCACCGCAACAATGTGTGCTGGAGACACTCGAGCGGGGCCCGAGCAAACT 735
Qy      890 TGCACGAGCGCTGCCAGGCGGATTCGGAGAGCCCTCTGCTGTGCTGAACGATGGCGCA 949
Db      736 TGCACGAGCGCTGCCAGGCGGATTCGGAGAGCCCTCTGCTGTGCTGAACGATGGCGCA 795
Qy      950 TGA-CTTTGGTGGGATCATCAGCTGGGG-CCTGGGCTGTGGACAGAGGATGCCGGG 1007
Db      796 TGACCTTTGGTGGGATCATCAGCTGGGGCCCTGGGCTGTTCACAGAGGATGTTCCGG 855
Qy      1008 TGTGTACACAAGGTTACCACTACCTAGACTGATTCGTGACACA 1054
Db      856 TGGTGTACCAAGTTTACCACTACCTAAACCGGATTCGGGACACA 902

```

```

RESULT 8
LOCUS   BUB45014
DEFINITION BUB45014.1 330 bp mRNA linear EST 16-OCT-2002
IMAGE:6578533 5', mRNA sequence.
ACCESSION BUB45014
VERSION   BUB45014.1 GI:24029455
KEYWORDS EST.
SOURCE   Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Contact: Robert Strausberg, Ph.D.
          Email: cgapbs-r@mail.nih.gov
          Tissue Procurement: ATCC
          cDNA Library Preparation: Rubin Laboratory
          DNA Sequencing by: Agencourt Bioscience Corporation
          Cloning Distribution: MGC clone distribution information can be
          found through the I.M.A.G.E. Consortium/LLNL at:
          http://image.llnl.gov
          Plate: LLCM2780 row: 9 column: 13
          High quality sequence stop: 639.
          Location/Qualifiers
            1. 930
              /organism="Homo sapiens"
              /mol_type="mRNA"
              /db_xref="taxon:9606"
              /clone="IMAGE:6578533"
              /tissue_type="teratocarcinoma, cell line"
              /lab_host="DH10B (phage-resistant)"
              /clone_lib="NIH_MGC_109"
              /notes="Organ: ovary; Vector: pOTB7; Site: 1: EcoRI; Site: 2:
              XhoI; cDNA made by oligo-dT priming. Directionally cloned
              into EcoRI/XhoI sites using the following 5' adaptor:
              GGCACGAG(G). Library constructed by Ling Hong in the
              laboratory of Gerald M. Rubin (University of California,
              Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
              Superscript II RT (Life Technologies). Note: this is a
              NIH_MGC Library."

```

BASE COUNT 199 a 279 c 266 g 186 t

# ORIGIN

```

Query Match 74.8%; Score 796.2; DB 13; Length 930;
Best Local Similarity 99.4%; Pred. No. 3.8e-197;
Matches 820; Conservative 0; Mismatches 3; Indels 2; Gaps 2;

Qy      234 GACGTGGAGTACTGTGATGTGCTCTCTGCTCCACCTGCGGCTTGAGACAGTACAGCA 293
Db      1 GACGTGGAGTACTGTGATGTGCTCTCTGCTCCACCTGCGGCTTGAGACAGTACAGCA 60
Qy      294 GCCTCAGTTTCGATCAAAAGAGGCTCTTCGCCGACATCGCCTCCACCCCTGCGAGGC 353
Db      61 GCCTCAGTTTCGATCAAAAGAGGCTCTTCGCCGACATCGCCTCCACCCCTGCGAGGC 120
Qy      354 TGCCATCTTTGCCAAGCACAGGAGGTGCCCGGAGAGCGGTTCTGTGCGGGGCATACT 413
Db      121 TGCCATCTTTGCCAAGCACAGGAGGTGCCCGGAGAGCGGTTCTGTGCGGGGCATACT 180
Qy      414 CATCAGCTCTCTGTGGATTTCTCTGCGGCCACTGCTTCCAGGAGAGGTTTCCGCCCA 473
Db      181 CATCAGCTCTCTGTGGATTTCTCTGCGGCCACTGCTTCCAGGAGAGGTTTCCGCCCA 240
Qy      474 CCACCTCAGCGTGATCTTTGGGCGAGAACATACCGGGTGTCTCCGCGAGGAGGAGCA 533
Db      241 CCACCTCAGCGTGATCTTTGGGCGAGAACATACCGGGTGTCTCCGCGAGGAGGAGCA 300
Qy      534 ATTTGAAGTCGAAAAATACATTTGTCCATAAGGAATTCGATGATGACACTTACGACAATGA 593
Db      301 ATTTGAAGTCGAAAAATACATTTGTCCATAAGGAATTCGATGATGACACTTACGACAATGA 360
Qy      594 CATTGCGCTGTGAGCTGAAATCGGATTCGTCCTGCTGCTGCGGCTGCGGAGAGCAGCTGGT 653
Db      361 CATTGCGCTGTGAGCTGAAATCGGATTCGTCCTGCTGCTGCGGCTGCGGAGAGCAGCTGGT 420
Qy      654 CCGCACTGTGCTTCCCGCGGACCTGACCTGCGGACCTGCGGACCTGCGGACCTGCGGAC 480
Db      421 CCGCACTGTGCTTCCCGCGGACCTGACCTGCGGACCTGCGGACCTGCGGACCTGCGGAC 440
Qy      714 CTCGGCTACGCGAAGCATGAGCGCTTGTCTCTCTTCTATTCGAGCGGCTGAAAGAGGC 773
Db      481 CTCGGCTACGCGAAGCATGAGCGCTTGTCTCTCTTCTATTCGAGCGGCTGAAAGAGGC 540
Qy      774 TCATGTGAGCTGTGACCCATCCAGCGCTGACATCAACATTTACTTTAACAGAACAGT 833
Db      541 TCATGTGAGCTGTGACCCATCCAGCGCTGACATCAACATTTACTTTAACAGAACAGT 600
Qy      834 CACCGACAATGCTGTGCTGAGAGACATCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGG 893
Db      601 CACCGACAATGCTGTGCTGAGAGACATCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGG 660
Qy      894 CGAGCGCTGCGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 953
Db      661 CGAGCGCTGCGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 720
Qy      954 TTTGGTGGGATCATCAGCTGGGGCTGGGCTGTGGACAGAGGATGTCCC-GGGTGTGT 1012
Db      721 TTTGGTGGGATCATCAGCTGGGGCTGGGCTGTGGACAGAGGATGTCCC-GGGTGTGT 780
Qy      1013 ACA-CAAAAGTTACCAACTACCTAGACTGATTCGATGATTCGATGATTCGATGATTCG 1056
Db      781 ACACCCAAAGGTTACCAACTACCTAGACTGATTCGATGATTCGATGATTCGATGATTCG 825

```

```

RESULT 9
LOCUS   BUI57720
DEFINITION BUI57720.1 888 bp mRNA linear EST 04-SEP-2002
IMAGE:8043192 NIH_MGC_110 Homo sapiens cDNA clone IMAGE:6084476
          5', mRNA sequence.
ACCESSION BUI57720
VERSION   BUI57720.1 GI:22671252
KEYWORDS EST.
SOURCE   Homo sapiens (human)
ORGANISM Homo sapiens

```



Gruber (Invitrogen). Research Genetics tracking code  
023. Note: this is a NIH MGC Library."

BASE COUNT 204 a 272 c 256 g 184 t

ORIGIN

Query Match 72.1%; Score 767.4; DB 12; Length 916;  
Best Local Similarity 99.7%; Pred. No. 1.3e-189;  
Matches 779; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 286 TACAGCCAGCCTCAGTTTCGATCAAGAGAGGGCTCTTCGCGACATCGCTCCACACCC 345  
DB 1 TACAGCCAGCCTCAGTTTCGATCAAGAGAGGGCTCTTCGCGACATCGCTCCACACCC 60

QY 346 TGGCAGGCTGCCATCTTTGCCAGACAGGAGGTGCCGAGAGCGGTTCTGTGCGGG 405  
DB 61 TGGCAGGCTGCCATCTTTGCCAGACAGGAGGTGCCGAGAGCGGTTCTGTGCGGG 120

QY 406 GGCATACATCATCAGCTCCTGCTGGATTCTCTGCGCGCCCACTGCTCCAGGAGAGGTTT 465  
DB 121 GGCATACATCATCAGCTCCTGCTGGATTCTCTGCGCGCCCACTGCTCCAGGAGAGGTTT 180

QY 466 CGCGCCCAACCTGACGGTGATCTTGGCGAGAACATACCGGGTGGTCCCTGGCGAGGAG 525  
DB 181 CGCGCCCAACCTGACGGTGATCTTGGCGAGAACATACCGGGTGGTCCCTGGCGAGGAG 240

QY 526 GAGCAGAAATTTGAAGTCGAAATAATACATTGTTCATTAAGAAATTCATGATGACATTAC 585  
DB 241 GAGCAGAAATTTGAAGTCGAAATAATACATTGTTCATTAAGAAATTCATGATGACATTAC 300

QY 586 GACATGACATTCGGCTGCTCAGCTGAATTCGGAATTCGTCGCGCTGCGCCAGGAGGAGC 645  
DB 301 GACATGACATTCGGCTGCTCAGCTGAATTCGGAATTCGTCGCGCTGCGCCAGGAGGAGC 360

QY 646 AGCGTGCTCCGACTGTGTGCTTCCCGCGGACCTCGAGCTGCCGGAATCGACAGGAG 705  
DB 361 AGCGTGCTCCGACTGTGTGCTTCCCGCGGACCTCGAGCTGCCGGAATCGACAGGAG 420

QY 706 TGTGAGCTCTCCGGCTACGGCAAGCATGAGCGCTTGTCTCTTCTATTCGGAGCGGCTG 765  
DB 421 TGTGAGCTCTCCGGCTACGGCAAGCATGAGCGCTTGTCTCTTCTATTCGGAGCGGCTG 480

QY 766 AAGGAGGCTCATGTGACAGCTGTACCCATCCAGCGCTGCACATCACACATTTACTTAAC 825  
DB 481 AAGGAGGCTCATGTGACAGCTGTACCCATCCAGCGCTGCACATCACACATTTACTTAAC 540

QY 826 AGAAGAGTCACCGACAAATGCTGTGTGCTGAGACACTCGGAGCGGCGGCCCCAGGCA 885  
DB 541 AGAAGAGTCACCGACAAATGCTGTGTGCTGAGACACTCGGAGCGGCGGCCCCAGGCA 600

QY 886 AACTTGACAGCGCTGCGAGGCGATTCGGAGGCGCCCTGGTGTGTCTGAACGATGGC 945  
DB 601 AACTTGACAGCGCTGCGAGGCGATTCGGAGGCGCCCTGGTGTGTCTGAACGATGGC 660

QY 946 CGCATGCTTTGTGGGATCATCAGCTGGGGCTGGGCTGTGACAGAGAGATGTCCTCG 1005  
DB 661 CGCATGCTTTGTGGGATCATCAGCTGGGGCTGGGCTGTGACAGAGAGATGTCCTCG 720

QY 1006 GGTGTGTACAAAGGTTTACCACTACCTAGAC - TGGATTGCTGACAACTGCGACCGTG 1064  
DB 721 GGTGTGTACAAAGGTTTACCACTACCTAGACCTTGGATTGCTGACAACTGCGACCGTG 780

QY 1065 A 1065  
DB 781 A 781

RESULT 11  
BQ927899  
LOCUS  
DEFINITION BQ927899 950 bp mRNA linear EST 20-AUG-2002  
5', mRNA sequence.  
ACCESSION BQ927899  
VERSION BQ927899.1 GI:22342930

KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE 1 (bases 1 to 950)  
AUTHORS NIH-MGC http://mgi.nci.nih.gov/  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: csapbe@mail.nih.gov  
Tissue Procurement: DCTD/DTP/Gazdar  
cDNA Library Preparation: Rubin Laboratory  
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LUCM2550 row: p column: 02  
High quality sequence start: 35  
High quality sequence stop: 684.  
Location/Qualifiers  
1. 950  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clones="IMAGE:6373489"  
/tissue types="large cell carcinoma"  
/lab host="DH10B (phage-resistant)"  
/clone\_lib="NIH\_MGC\_18"  
/notes="Organ: lung; Vector: pOTB7; Site: 1: XhoI; Site: 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGACAGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH MGC Library."

BASE COUNT 207 a 283 c 279 g 181 t

ORIGIN

Query Match 71.8%; Score 765.2; DB 13; Length 950;  
Best Local Similarity 98.1%; Pred. No. 4.9e-189;  
Matches 806; Conservative 0; Mismatches 13; Indels 3; Gaps 3;

QY 210 CCAGCTGCTCAAGAAACGAGGCTGAGCTGGAGTACTGTGATGTCCTCTCTCTCCAC 269  
DB 43 CCGGCTGCTGAGAACCGAGGCTGAGCTGGAGTACTGTGATGTCCTCTCTCTCCAC 102

QY 270 CTGCGGCTGAGACAGTACAGCCCTCAGTTTTCGCATCAAGAGGAGGCTCTTCGCCGA 329  
DB 103 CTGCGGCTGAGACAGTACAGCCCTCAGTTTTCGCATCAAGAGGAGGCTCTTCGCCGA 162

QY 330 CATGCGCTCCACCCCTGGCAGGCTGCATCTTTGCCAAGACAGAGAGTGCCTCCGAGA 389  
DB 163 CATGCGCTCCACCCCTGGCAGGCTGCATCTTTGCCAAGACAGAGAGTGCCTCCGAGA 222

QY 390 GCGTTTCTGTCGGGGGACATCTCATCAGCTCTGCTGATTTCTCTCTGCGGCCACTG 449  
DB 223 GCGTTTCTGTCGGGGGACATCTCATCAGCTCTGCTGATTTCTCTCTGCGGCCACTG 282

QY 450 CTTTCCAGGAGG--TTTCCGCCCCACCCCTGACGGTGTCTTTGGGAGAACATACATTCG 508  
DB 283 CTTTCCAGGAGGTTTCTGCCCCACCCCTGACGGTGTCTTTGGGAGAACATACATTCG 342

QY 509 TGTGCTCTGCGAGGAGGAGAGAAATTTGAAGTCGAAATAATCATTTGTCATAAGAAAT 569  
DB 343 TGTGCTCTGCGAGGAGGAGAGAAATTTGAAGTCGAAATAATCATTTGTCATAAGAAAT 402

QY 569 TCGATGATGACATTCAGCAATGACATTCGCGCTGCTGAGCTGCAATTCGATTCGTC 628  
DB 403 TCGATGATGACATTCAGCAATGACATTCGCGCTGCTGAGCTGCAATTCGATTCGTC 462



mRNA sequence.  
 BG763582 1 GI:14074235  
 EST.  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Homo sapiens (human)  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 NIH-MGC http://mgi.nci.nih.gov/.  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgapbs-remail.nih.gov  
 Tissue Procurement: ATCC/DCTD/DTF  
 cDNA Library Preparation: Ling Hong/Rubin Laboratory  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: LLCMI1719 row: n column: 21  
 High quality sequence stop: 810.  
 Location/Qualifiers  
 1..958  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:4861268"  
 /tissue\_type="melanotic melanoma, high MDR (cell line)"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="NIH MGC\_49"  
 /note="Organ: skin; Vector: pOTB7; Site 1: XhoI; Site 2:  
 EcoRI; cDNA made by oligo-dT priming. Directionally cloned  
 into EcoRI/XhoI sites using the following 5' adaptor:  
 GGCACGAG(G). Size-selected >500bp for average insert size  
 of Gerald M. Rubin (University of California, Berkeley)  
 using ZAP-cDNA synthesis kit (Stratagene) and Superscript  
 II RT (Life Technologies). Note: this is a NIH\_MGC  
 Library."

BASE COUNT 198 a 285 c 282 g 193 t  
 ORIGIN  
 Query Match 65.5%; Score 697.4; DB 12; Length 958;  
 Best Local Similarity 94.2%; Pred. No. 2.6e-171;  
 Matches 834; Conservative 0; Mismatches 36; Indels 15; Gaps 10;

Qy 1 TCTGAGGAAACAGTACTGCTTTGGGAAATGGGTGAGCTACCGTGGCAGCAGCAGC 60  
 Db 72 TCTGAGGAAACAGTACTGCTTTGGGAAATGGGTGAGCTACCGTGGCAGCAGCAGC 131  
 Qy 61 CTACCCGAGTCGGTCTCTGCTCCGCTGGAATCCATGATCCCTGATAGGCAAGGTT 120  
 Db 132 CTACCCGAGTCGGTCTCTGCTCCGCTGGAATCCATGATCCCTGATAGGCAAGGTT 191  
 Qy 121 TACACAGACAGAACCCAGTCCGAGCAGTGGGCTGGGCAACATATTAATCTGCCGG 180  
 Db 192 TACACAGACAGAACCCAGTCCGAGCAGTGGGCTGGGCAACATATTAATCTGCCGG 251  
 Qy 181 AATCTGATGGGATGCCAAGCCCTGTTGTCACCTGCTGAAGAACCCGAGGCTGACCTGG 240  
 Db 252 AATCTGATGGGATGCCAAGCCCTGTTGTCACCTGCTGAAGAACCCGAGGCTGACCTGG 311  
 Qy 241 GAGTACTGTGATGTCCTCTGCTCCAGCTGGGCTGAGACAGTACAGCCGCTCAG 300  
 Db 312 GAGTACTGTGATGTCCTCTGCTCCAGCTGGGCTGAGACAGTACAGCCGCTCAG 371  
 Qy 301 TTTCGCATCAAGAGGGGCTTTCCGCGACATCCCTCCACCCCTGGCAGGCTGCCATC 360  
 Db 372 TTTCGCATCAAGAGGGGCTTTCCGCGACATCCCTCCACCCCTGGCAGGCTGCCATC 431  
 Qy 361 TTTCGCCAAGCAGGAGGTCGCCGAGAGCGGTTCTCTGTGCGGGGGGCATCTACTATCAGC 420

Db 432 TTTCGCCAAGCAGGAGGTCGCCGAGAGCGGTTCTCTGTGCGGGGGGCATCTACTATCAGC 491  
 Qy 421 TCCTGTCTGAGTTCTCTGCGGCCACTGCTTCAGGAGAGGTTTCCGCCCCACACCTG 480  
 Db 492 TCCTGTCTGAGTTCTCTGCGGCCACTGCTTCAGGAGAGGTTTCTTCCGCCACACCTG 551  
 Qy 481 ACGTGTATCTTTGGGCAAGACATACCGGGTGTCTCCCTGGCGAGGAGGAGAAATTTGAA 540  
 Db 552 ACGTGTATCTTTGGGCAAGACATACCGGGTGTCTCCCTGGCGAGGAGGAGAAATTTGAA 611  
 Qy 541 GTCGAAAATACATTGTCCATTAAGGAATTCGATGATGACACTTACGACAAATGCAATTCGCG 600  
 Db 612 GTCGAAAATACATTGTCCATTAAGGAATTCGATGATGACACTTACGACAAATGCAATTCGCG 671  
 Qy 601 CTGCTGAGCTGAAATCGGATTCGT-CCCGCTGTGCGCCAGGAGAGCAGC-GTGCTCCGCA 658  
 Db 672 CTGCTGAGCTGAAATCGGATTCGTCCCGCTGTGCGCCAGGAGAGCAGC-GTGCTCCGCA 731  
 Qy 659 CTGCTGCTCTT-CCCGCGGCGACCTGCGAGCTGCGGACTGGAC-GGAGTGTGAGCTCT 715  
 Db 732 CTGCTGCTCTTCCCGCGGCGACCTGCGAGCTGCGGACTGGAGCGGAGTGTGAGCTCT 791  
 Qy 716 CCG-GCTACCGCAAGCAT--GAGGCTTGTCTCTTCTTATTCGAGCGGCTGAGAGG 772  
 Db 792 CCGAGCTACCGCAAGCATTCGAGGCGCTGCTCTCTTCTTATTCGAGCGG-TGAAGGAGG 850  
 Qy 773 CTATGTCAGACTGTATCCCATCCAGCGCTGCGACATCACAACATTTACTTAACAGAAACAG 832  
 Db 851 CTATG--CGACTGGACCCATTCACCGGTGGCCATTCACAACATTTACTT-ACAGAAACAG 907  
 Qy 833 TCACCGACAAACATCTGTGTGCTGGGAGACACTCGGAGCGGCGGCG 877  
 Db 908 TACC---GAAATGGTGTGTGCGGACCCCGACGGGGGCCGAGC 949

BO960216 974 bp mRNA linear EST 21-AUG-2002  
 AGENCOURT 8945604 NIH\_MGC\_101 Homo sapiens cDNA clone IMAGE:6459629  
 5', mRNA sequence.  
 BO960216  
 BO960216.1 GI:22375694  
 EST.  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Homo sapiens (human)  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 NIH-MGC http://mgi.nci.nih.gov/.  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgapbs-remail.nih.gov  
 Tissue Procurement: ATCC  
 cDNA Library Preparation: Rubin Laboratory  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: LLCM2645 row: a column: 06  
 High quality sequence stop: 603.  
 Location/Qualifiers  
 1..974  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:6459629"  
 /tissue\_type="epidermoid carcinoma, cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="NIH MGC\_101"  
 /note="Organ: lung; Vector: pOTB7; Site 1: EcoRI; Site 2:

RESULT 14  
 BO960216  
 LOCUS  
 DEFINITION  
 ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 COMMENT

FEATURES  
 source







```
Qy 617 CGGATTCGTCGGCTGTGCGCAGGAGAGACGCGTGGTCCGCACTGTGTGCCTTCCCCCGG 676
Db |||||
Qy 421 CGGATTCGTCGGCTGTGCGCAGGAGAGACGCGTGGTCCGCACTGTGTGCCTTCCCCCGG 480
Db |||||
Qy 677 CGGACCTGCAGTCGCCGACTGGACGGAGTGTGAGCTCTCGGCTACGGCAAGCATGAGG 736
Db |||||
Qy 481 CGGACCTGCAGTCGCCGACTGGACGGAGTGTGAGCTCTCGGCTACGGCAAGCATGAGG 540
Db |||||
Qy 737 CCTTGTCTCCTTTCTATTTCGGAGCGGCTGAAGGAGGCTCATGTCAAGACTGTACCCATCCA 796
Db |||||
Qy 541 CCTTGTCCCTTTCTATTTCGGAGCGGCTGAAGGAGGCTCATGTCAAGACTGTACCCATCCA 600
Db |||||
Qy 797 GCCGCTGCACATCACAACTTTACTTAAACAGAAAGTCAACGCAACATGCTGTGTGCTG 856
Db |||||
Qy 601 GCCGCTGCACATCACAACTTTACTTAAACAGAAAGTCAACGCAACATGCTGTGTGCTG 660
Db |||||
Qy 857 GAGACACTCGGAGCGGGCGGCCAGGCAAACTTGCAGAGCGCTGC 903
Db |||||
Qy 661 GAGACACTCGGAGCGGGCGGG-CCGAGGCAAACTTGCAGAGCGCTGC 706
Db |||||
```

Search completed: August 19, 2003, 08:29:20  
Job time : 3116.64 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 18, 2003, 19:53:48 ; Search time 4471.78 Seconds  
(without alignments)  
9743.047 Million cell updates/sec

Title: US-09-987-455-4  
Perfect score: 1065  
Sequence: 1 tctgagggaacacgtgactg.....gtgacaacatgcgaccgtga 1065  
Scoring table: IDENTITY\_NUC  
Gapop 10.0, Gapext 1.0

Searched: 2888711 seqs, 20454813386 residues  
Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl.\*

- 1: gb\_ba.\*
- 2: gb\_htg.\*
- 3: gb\_in.\*
- 4: gb\_om.\*
- 5: gb\_ov.\*
- 6: gb\_pat.\*
- 7: gb\_ph.\*
- 8: gb\_pl.\*
- 9: gb\_pr.\*
- 10: gb\_ro.\*
- 11: gb\_sts.\*
- 12: gb\_sy.\*
- 13: gb\_un.\*
- 14: gb\_vl.\*
- 15: em\_ba.\*
- 16: em\_fun.\*
- 17: em\_hum.\*
- 18: em\_in.\*
- 19: em\_mu.\*
- 20: em\_om.\*
- 21: em\_or.\*
- 22: em\_ov.\*
- 23: em\_pat.\*
- 24: em\_ph.\*
- 25: em\_pl.\*
- 26: em\_ro.\*
- 27: em\_sts.\*
- 28: em\_un.\*
- 29: em\_vl.\*
- 30: em\_htg\_hum.\*
- 31: em\_htg\_inv.\*
- 32: em\_htg\_other.\*
- 33: em\_htg\_mus.\*
- 34: em\_htg\_pln.\*
- 35: em\_htg\_rtd.\*
- 36: em\_htg\_mam.\*
- 37: em\_htg\_vrt.\*
- 38: em\_sy.\*
- 39: em\_htgo\_hum.\*
- 40: em\_htgo\_mus.\*
- 41: em\_htgo\_other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1065	100.0	1065	6	AX431270	AX431270 Sequence
2	1065	100.0	1065	6	AX431273	AX431273 Sequence
3	1065	100.0	1068	6	A27431	A27431 DNA sequence
4	1065	100.0	1068	6	AR059987	AR059987 Sequence
5	1065	100.0	1068	6	E01935	E01935 Synthetic D
6	1065	100.0	1170	6	A27727	A27727 DNA sequence
7	1065	100.0	1170	6	E01945	E01945 Synthetic D
8	1065	100.0	1314	6	A27435	A27435 DNA sequence
9	1065	100.0	1314	6	E01937	E01937 Synthetic D
10	1065	100.0	1419	6	A27725	A27725 DNA sequence
11	1065	100.0	1419	6	AR059996	AR059996 Sequence
12	1065	100.0	1419	6	E01944	E01944 Synthetic D
13	1065	100.0	1689	6	I06609	I06609 Sequence 44
14	1065	100.0	1689	6	I06614	I06614 Sequence 46
15	1065	100.0	1689	6	I08789	I08789 Sequence 3
16	1065	100.0	1780	6	A06611	A06611 Synthetic n
17	1065	100.0	1780	6	A10226	A10226 Synthetic D
18	1065	100.0	1836	6	E01176	E01176 DNA encodin
19	1065	100.0	1836	6	I01581	I01581 Sequence 1
20	1065	100.0	1981	6	I01583	I01583 Sequence 3
21	1065	100.0	1989	6	I01256	I01256 Sequence 2
22	1065	100.0	2099	6	A07207	A07207 Artificial
23	1065	100.0	2100	6	A07197	A07197 H.sapiens m
24	1065	100.0	2100	6	A27429	A27429 cDNA sequen
25	1065	100.0	2100	6	E01934	E01934 DNA encodin
26	1065	100.0	2101	6	AR059986	AR059986 Sequence
27	1065	100.0	2151	6	E00654	E00654 cDNA encodi
28	1065	100.0	2162	6	AR017909	AR017909 Sequence
29	1065	100.0	2162	6	AR017931	AR017931 Sequence
30	1065	100.0	2162	6	I18796	I18796 Sequence 3
31	1065	100.0	2249	6	A07209	A07209 Artificial
32	1065	100.0	2544	6	I08530	I08530 Sequence 1
33	1065	100.0	2547	6	E01163	E01163 cDNA encodi
34	1065	100.0	2547	6	I07841	I07841 Sequence 1
35	1065	100.0	2547	6	I07989	I07989 Sequence 1
36	1065	100.0	2549	6	I09622	I09622 Sequence 1
37	1065	100.0	2549	6	E01055	E01055 DNA encodin
38	1065	100.0	2560	6	A01465	A01465 H.sapiens t
39	1065	100.0	7533	6	I08638	I08638 Sequence 2
40	1064	99.9	1689	6	I08637	I08637 Sequence 1
41	1063.4	99.8	1170	6	AR059997	AR059997 Sequence
42	1063.4	99.8	1314	6	AR059989	AR059989 Sequence
43	1063.4	99.8	1506	6	E02247	E02247 cDNA sequen
44	1063.4	99.8	1680	6	I06422	I06422 Sequence 40
45	1063.4	99.8	1689	6	A14927	A14927 tPA-DNA fro

ALIGNMENTS

RESULT 1	AX431270	AX431270	Sequence 4 from Patent WO0240650.	1065 bp	DNA	linear	PAT 28-JUN-2002
LOCUS	AX431270	Sequence 4 from Patent WO0240650.					
DEFINITION	AX431270	Sequence 4 from Patent WO0240650.					
ACCESSION	AX431270	Sequence 4 from Patent WO0240650.					
VERSION	AX431270.1	GI:21656152					
KEYWORDS		synthetic construct					
SOURCE		synthetic construct					
ORGANISM		artificial sequences.					
REFERENCE		1					
AUTHORS		Goetz,F., Werner,R.G., Tayapiwatana,C., Manosroi,A. and Manosroi,J.					
TITLE		Methods for large scale production of recombinant dna-derived tpa					
JOURNAL		or k28 molecules					
		Patent: WO 0240650-A 4 23-MAY-2002;					

BOEHRINGER INGELHEIM INT (DE)	
FEATURES	Location/Qualifiers
source	1..1065
	/organism="synthetic construct"
	/mol_type="genomic DNA"
	/db_xref="taxon:32630"
	/note="coding sequence for K2S protein"
BASE COUNT	225 a 314 c 312 g 214 t
ORIGIN	
Query Match 100.0%; Score 1065; DB 6; Length 1065;	
Best Local Similarity 100.0%; Pred. No. 4.7e-248;	
Matches 1065; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
Qy	1 TCTGAGGGAACAGTACTGCTTCTGGGATGGGTGAGCTTACCGTGGGACGACAGC 60
Db	1 TCTGAGGGAACAGTACTGCTTCTGGGATGGGTGAGCTTACCGTGGGACGACAGC 60
Qy	61 CTCACCGAGTGGGTGCTTCTGCTCCGCTGGGAATTCATGATCCTGATAGGCAAGTT 120
Db	61 CTCACCGAGTGGGTGCTTCTGCTCCGCTGGGAATTCATGATCCTGATAGGCAAGTT 120
Qy	121 TACACAGACAGAACCCAGTGGCCAGGACCTGGGCTGGGCAACATAATTAATCTGCCGG 180
Db	121 TACACAGACAGAACCCAGTGGCCAGGACCTGGGCTGGGCAACATAATTAATCTGCCGG 180
Qy	181 AATCTGATGGGATGCCAAGCCCTGGTGGCCAGCTGTGCAAGAACCCGAGCTCACGTGG 240
Db	181 AATCTGATGGGATGCCAAGCCCTGGTGGCCAGCTGTGCAAGAACCCGAGCTCACGTGG 240
Qy	241 GAGTACTGTGATGTGCCCTCTCTGCTCCACCTGGGGCTTGAGACAGTACAGCCCTCAG 300
Db	241 GAGTACTGTGATGTGCCCTCTCTGCTCCACCTGGGGCTTGAGACAGTACAGCCCTCAG 300
Qy	301 TTTGCGATCAAGAGGGCTCTTGGCCGACATGCTCCACCCCTGGGAGGCTGCCATC 360
Db	301 TTTGCGATCAAGAGGGCTCTTGGCCGACATGCTCCACCCCTGGGAGGCTGCCATC 360
Qy	361 TTTGCGAAGCAGAGGAGTGGCCGAGAGCGGTTCTGTGGGGGACATCTACTCATCAGC 420
Db	361 TTTGCGAAGCAGAGGAGTGGCCGAGAGCGGTTCTGTGGGGGACATCTACTCATCAGC 420
Qy	421 TCCTGTGTGATTTCTCTGTGGCCGCTCTGCTTCCAGGAGGTTTCCGCCCCACCATCTG 480
Db	421 TCCTGTGTGATTTCTCTGTGGCCGCTCTGCTTCCAGGAGGTTTCCGCCCCACCATCTG 480
Qy	481 ACGGTGATCTTGGGAGAAATACATCGAATTCGATGATGACCTTACGACAAATGCAATTCG 540
Db	481 ACGGTGATCTTGGGAGAAATACATCGAATTCGATGATGACCTTACGACAAATGCAATTCG 540
Qy	541 GTCGAAATACATTCGATGATGACCTTACGACAAATGCAATTCGATGATGACCTTACGACAA 600
Db	541 GTCGAAATACATTCGATGATGACCTTACGACAAATGCAATTCGATGATGACCTTACGACAA 600
Qy	601 CTGCTGACGTGAATTCGATGATGACCTTACGACAAATGCAATTCGATGATGACCTTACGAC 660
Db	601 CTGCTGACGTGAATTCGATGATGACCTTACGACAAATGCAATTCGATGATGACCTTACGAC 660
Qy	661 GTGTGCTTCCCGCGGAGACTGTCAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAG 720
Db	661 GTGTGCTTCCCGCGGAGACTGTCAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAG 720
Qy	721 TACGCAAGCATGAGGCTTCTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 780
Db	721 TACGCAAGCATGAGGCTTCTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 780
Qy	781 AGACTGTACCATTCAGCGCTGACATCAACAACTTACTTAAACAGAACTGACCGAC 840
Db	781 AGACTGTACCATTCAGCGCTGACATCAACAACTTACTTAAACAGAACTGACCGAC 840
Qy	841 AACATGTGTGTGTGGAGACATCTGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 900
Db	841 AACATGTGTGTGTGGAGACATCTGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 900
RESULT 2	
AX431273	
LOCUS	
DEFINITION	
ACCESSION	
VERSION	
KEYWORDS	
SOURCE	
ORGANISM	
REFERENCE	
AUTHORS	
TITLE	
JOURNAL	
FEATURES	
source	
1..1065	
/organism="synthetic construct"	
/mol_type="genomic DNA"	
/db_xref="taxon:32630"	
/note="coding sequence for K2S protein"	
BASE COUNT	225 a 314 c 312 g 214 t
ORIGIN	
Query Match 100.0%; Score 1065; DB 6; Length 1065;	
Best Local Similarity 100.0%; Pred. No. 4.7e-248;	
Matches 1065; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
Qy	1 TCTGAGGGAACAGTACTGCTTCTGGGATGGGTGAGCTTACCGTGGGACGACAGC 60
Db	1 TCTGAGGGAACAGTACTGCTTCTGGGATGGGTGAGCTTACCGTGGGACGACAGC 60
Qy	61 CTCACCGAGTGGGTGCTTCTGCTCCGCTGGGAATTCATGATCCTGATAGGCAAGTT 120
Db	61 CTCACCGAGTGGGTGCTTCTGCTCCGCTGGGAATTCATGATCCTGATAGGCAAGTT 120
Qy	121 TACACAGACAGAACCCAGTGGCCAGGACCTGGGCTGGGCAACATAATTAATCTGCCGG 180
Db	121 TACACAGACAGAACCCAGTGGCCAGGACCTGGGCTGGGCAACATAATTAATCTGCCGG 180
Qy	181 AATCTGATGGGATGCCAAGCCCTGGTGGCCAGCTGTGCAAGAACCCGAGCTCACGTGG 240
Db	181 AATCTGATGGGATGCCAAGCCCTGGTGGCCAGCTGTGCAAGAACCCGAGCTCACGTGG 240
Qy	241 GAGTACTGTGATGTGCCCTCTCTGCTCCACCTGGGGCTTGAGACAGTACAGCCCTCAG 300
Db	241 GAGTACTGTGATGTGCCCTCTCTGCTCCACCTGGGGCTTGAGACAGTACAGCCCTCAG 300
Qy	301 TTTGCGATCAAGAGGGCTCTTGGCCGACATGCTCCACCCCTGGGAGGCTGCCATC 360
Db	301 TTTGCGATCAAGAGGGCTCTTGGCCGACATGCTCCACCCCTGGGAGGCTGCCATC 360
Qy	361 TTTGCGAAGCAGAGGAGTGGCCGAGAGCGGTTCTGTGGGGGACATCTACTCATCAGC 420
Db	361 TTTGCGAAGCAGAGGAGTGGCCGAGAGCGGTTCTGTGGGGGACATCTACTCATCAGC 420
Qy	421 TCCTGTGTGATTTCTCTGTGGCCGCTCTGCTTCCAGGAGGTTTCCGCCCCACCATCTG 480
Db	421 TCCTGTGTGATTTCTCTGTGGCCGCTCTGCTTCCAGGAGGTTTCCGCCCCACCATCTG 480
Qy	481 ACGGTGATCTTGGGAGAAATACATCGAATTCGATGATGACCTTACGACAAATGCAATTCG 540
Db	481 ACGGTGATCTTGGGAGAAATACATCGAATTCGATGATGACCTTACGACAAATGCAATTCG 540
Qy	541 GTCGAAATACATTCGATGATGACCTTACGACAAATGCAATTCGATGATGACCTTACGACAA 600
Db	541 GTCGAAATACATTCGATGATGACCTTACGACAAATGCAATTCGATGATGACCTTACGACAA 600
Qy	601 CTGCTGACGTGAATTCGATGATGACCTTACGACAAATGCAATTCGATGATGACCTTACGAC 660
Db	601 CTGCTGACGTGAATTCGATGATGACCTTACGACAAATGCAATTCGATGATGACCTTACGAC 660
Qy	661 GTGTGCTTCCCGCGGAGACTGTCAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAG 720
Db	661 GTGTGCTTCCCGCGGAGACTGTCAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAG 720
Qy	721 TACGCAAGCATGAGGCTTCTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 780
Db	721 TACGCAAGCATGAGGCTTCTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 780
Qy	781 AGACTGTACCATTCAGCGCTGACATCAACAACTTACTTAAACAGAACTGACCGAC 840
Db	781 AGACTGTACCATTCAGCGCTGACATCAACAACTTACTTAAACAGAACTGACCGAC 840
Qy	841 AACATGTGTGTGTGGAGACATCTGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 900
Db	841 AACATGTGTGTGTGGAGACATCTGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 900

```
Db 421 TCCTGCTGGATTCTCTCTGCCGCCACACCTGCTTTCAGGAGAGGTTTCCGCCCCACACCTG 480
Qy 481 ACGGTGATCTTTGGGAGAACATACACGGGTGGTCCCTGCGAGGAGGAGACAGAAATTTGAA 540
Db 481 ACGGTGATCTTTGGGAGAACATACACGGGTGGTCCCTGCGAGGAGGAGACAGAAATTTGAA 540
Qy 541 GTCGAAAATACATTGTCTCAAGAAATTCGATGATGACACTTACGACAATGACATTTGCG 600
Db 541 GTCGAAAATACATTGTCTCAAGAAATTCGATGATGACACTTACGACAATGACATTTGCG 600
Qy 601 CTGCTGCAGCTGAATTCGGATTCGTCCTGCTGTGCCAGGAGAGCAGAGTGTCCGCACT 660
Db 601 CTGCTGCAGCTGAATTCGGATTCGTCCTGCTGTGCCAGGAGAGCAGAGTGTCCGCACT 660
Qy 661 GTGTGCTTCCCGCGGACCTGAGCTGCGGACTGAGCGAGTGTGAGCTCTCCGCG 720
Db 661 GTGTGCTTCCCGCGGACCTGAGCTGCGGACTGAGCGAGTGTGAGCTCTCCGCG 720
Qy 721 TACGCAAGCATGAGGCTTGTCTCTCTTCTTATTCGGAGCGGCTGAAGAGGCTCATGTC 780
Db 721 TACGCAAGCATGAGGCTTGTCTCTCTTCTTATTCGGAGCGGCTGAAGAGGCTCATGTC 780
Qy 781 AGACTGTATCCCATCCAGCGCTGCACATCAACAATTTACTTAACAGAACAGTCAACGAC 840
Db 781 AGACTGTATCCCATCCAGCGCTGCACATCAACAATTTACTTAACAGAACAGTCAACGAC 840
Qy 841 AACATGCTGTGTGCTGGAGACACTCGAGCGCGGCGGCCCGGCGGAACTTGCACGCGCC 900
Db 841 AACATGCTGTGTGCTGGAGACACTCGAGCGCGGCGGCCCGGCGGAACTTGCACGCGCC 900
Qy 901 TGCCAGGCGGATTCGGAGGCGCCCTGCTGTGTCTGAACGATGCGCGCATGATTTGCTG 960
Db 901 TGCCAGGCGGATTCGGAGGCGCCCTGCTGTGTCTGAACGATGCGCGCATGATTTGCTG 960
Qy 961 GGCATCATCAGCTGGGCGCTGGCTGTGCAGACAGAGGATGTCGCGGCTGTACACAAAG 1020
Db 961 GGCATCATCAGCTGGGCGCTGGCTGTGTGCAGACAGAGGATGTCGCGGCTGTACACAAAG 1020
Qy 1021 GTTACCACTACCTAGACTGGATTCTGTGACAAACATGCGACCGTGA 1065
Db 1021 GTTACCACTACCTAGACTGGATTCTGTGACAAACATGCGACCGTGA 1065

RESULT 3
LOCUS A27431 1068 bp DNA linear PAT 09-JUL-2002
DEFINITION DNA sequence of coding region in pTRKpAdeltatrp.
ACCESSION A27431
VERSION A27431.1 GI:21727228
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM artificial sequences.
REFERENCE 1 (bases 1 to 1068)
AUTHORS Kobayashi,M.
TITLE New tissue plasminogen activator
JOURNAL Patent: EP 0302456-A 31 08-FEB-1989;
FUJISAWA PHARMACEUTICAL CO., LTD
FEATURES
source
Location/Qualifiers
1..1068
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
BASE COUNT 226 a 314 c 313 g 215 t
ORIGIN
Query Match 100.0%; Score 1065; DB 6; Length 1068;
Best Local Similarity 100.0%; Pred. No. 4.7e-248;
Matches 1065; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 TCTGAGGAAACAGTACTGCTCTTTGGGAATGGGTACGCTTACCGTGGCAGCGACAGC 60
Db 1 TCTGAGGAAACAGTACTGCTCTTTGGGAATGGGTACGCTTACCGTGGCAGCGACAGC 60
```

RESULT 4  
AR059987

LOCUS AR059987 1068 bp DNA linear PAT 29-SEP-1999  
DEFINITION Sequence 44 from patent US 5840533.  
ACCESSION AR059987  
VERSION AR059987.1 GI:5986437  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 1068)  
AUTHORS Niwa,M., Saito,Y., Sasaki,H., Hayashi,M., Notani,J. and Kobayashi,M.  
TITLE Tissue plasminogen activator  
JOURNAL Patent: US 5840533-A 44 24-NOV-1998;  
FEATURES  
Location/Qualifiers  
source 1..1068  
BASE COUNT 226 a 314 c 313 g 215 t  
ORIGIN  
Query Match 100.0%; Score 1065; DB 6; Length 1068;  
Best Local Similarity 100.0%; Pred. No. 4.7e-248;  
Matches 1065; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 TCTGAGGAAACAGTGACTGCTACTTTGGGAATGGGTGAGCTACCGTGGCAACGACAGC 60  
DB 4 TCTGAGGAAACAGTGACTGCTACTTTGGGAATGGGTGAGCTACCGTGGCAACGACAGC 63  
QY 61 CTCACCGAGTCGGTGGCTCTCCCTCCGCTGGGAATTCATGATCCTCATAGGCAGGTT 120  
DB 64 CTCACCGAGTCGGTGGCTCTCCCTCCGCTGGGAATTCATGATCCTCATAGGCAGGTT 123  
QY 121 TACACAGACAGAACCCAGTGCACGACCTGGGCTGGGCAACATAATTAATCTGCCGG 180  
DB 124 TACACAGACAGAACCCAGTGCACGACCTGGGCTGGGCAACATAATTAATCTGCCGG 183  
QY 181 AATCTGTATGGGATGCCAAGCCCTGGTGCCACATGCTGCTGAAGAACCGCAGGCTGACGTGG 240  
DB 184 AATCTGTATGGGATGCCAAGCCCTGGTGCCACATGCTGCTGAAGAACCGCAGGCTGACGTGG 243  
QY 241 GAGTACTGTATGTCCTCTCTGCTCCACATGCTGCGGCTGAGACATGACGACGCTCAG 300  
DB 244 GAGTACTGTATGTCCTCTCTGCTCCACATGCTGCGGCTGAGACATGACGACGCTCAG 303  
QY 301 TTTCGCATCAAGAGGAGGCTCTTCGCCACATCGCTCCACCCCTGCGAGGCTGCCATC 360  
DB 304 TTTCGCATCAAGAGGAGGCTCTTCGCCACATCGCTCCACCCCTGCGAGGCTGCCATC 363  
QY 361 TTTCGCATCAAGAGGAGGCTCTTCGCCACATCGCTCCACCCCTGCGAGGCTGCCATC 420  
DB 364 TTTCGCATCAAGAGGAGGCTCTTCGCCACATCGCTCCACCCCTGCGAGGCTGCCATC 423  
QY 421 TCCTGCTGGATTCCTCTCGCCGCCACATGCTTCCAGGAGAGGTTTCGCCGCCACCATC 480  
DB 424 TCCTGCTGGATTCCTCTCGCCGCCACATGCTTCCAGGAGAGGTTTCGCCGCCACCATC 483  
QY 481 ACGGTGATCTTGGGAGAACATACCGGCTGGTCCCTGCGAGGAGGAGACAGAAATTTGAA 540  
DB 484 ACGGTGATCTTGGGAGAACATACCGGCTGGTCCCTGCGAGGAGGAGACAGAAATTTGAA 543  
QY 541 GTCGAAAATAATCATTTGTCATAAGGAATTCGATGATGACACTTACGACAAATGACATTCGC 600  
DB 544 GTCGAAAATAATCATTTGTCATAAGGAATTCGATGATGACACTTACGACAAATGACATTCGC 603  
QY 601 CTGCTGAGCTGAATTCGATTCCTCCGCTGTCGCCAGGAGAGGAGGCTGTCGCACT 660  
DB 604 CTGCTGAGCTGAATTCGATTCCTCCGCTGTCGCCAGGAGAGGAGGCTGTCGCACT 663  
QY 661 GTGTGCTTCCCGGGGAGACTGAGCTGCGGAGTGGAGTGGAGTCTCCGGC 720  
DB 664 GTGTGCTTCCCGGGGAGACTGAGCTGCGGAGTGGAGTGGAGTGGAGTCTCCGGC 723  
QY 721 TACGCAAGCATAGGCGCTTGTCTCTTTATTTCGGAGCGGCTGAAGGAGGCTCATGTC 780  
DB |||||

DB 724 TACGCAAGCATAGGCGCTTGTCTCTTTATTTCGGAGCGGCTGAAGGAGGCTCATGTC 783  
QY 781 AGACTGTACCATCCAGCCGCTGCACATCAACAATTTACTTACACAGACAGTCAACGAC 840  
DB 784 AGACTGTACCATCCAGCCGCTGCACATCAACAATTTACTTAAACAGAACAGTCAACGAC 843  
QY 841 AACATGCTGTGTCTGGAGACACTCGGAGCGCGGCGCCAGGCAAACTTTCACGACGCC 900  
DB 844 AACATGCTGTGTCTGGAGACACTCGGAGCGCGGCGCCAGGCAAACTTTCACGACGCC 903  
QY 901 TGCAGGGCGATTCCGAGGCGCCCTCGTGTGTCTGAACGATGCGCGCATGCTTTGGTG 960  
DB 904 TGCAGGGCGATTCCGAGGCGCCCTCGTGTGTCTGAACGATGCGCGCATGCTTTGGTG 963  
QY 961 GGCATCATCAGCTGGGCGCTGGGCTGGGACAGAGGATGTCCCGGTGTGTACACAAG 1020  
DB 964 GGCATCATCAGCTGGGCGCTGGGCTGGGACAGAGGATGTCCCGGTGTGTACACAAG 1023  
QY 1021 GTTACCAACTACCTAGACTGGATTCTGTGACAACTGCGACCGTGA 1065  
DB 1024 GTTACCAACTACCTAGACTGGATTCTGTGACAACTGCGACCGTGA 1068  
RESULT 5  
E01935  
LOCUS Synthetic DNA encoding new tissue plasminogen activator (t-PA).  
DEFINITION  
ACCESSION E01935  
VERSION E01935.1 GI:2170183  
KEYWORDS JP 1989104167-A/2.  
SOURCE synthetic construct  
ORGANISM synthetic construct  
REFERENCE 1 (bases 1 to 1068)  
AUTHORS Niwa,M., Saito,Y., Sasaki,H., Hayashi,M., Notani,J. and Kobayashi,M.  
TITLE NOVEL TISSUE PLASMINOGEN ACTIVATION FACTOR  
JOURNAL Patent: JP 1989104167-A 2 21-APR-1989;  
FUJISAWA PHARMACEUT CO LTD  
COMMENT OS Artificial gene  
OC Artificial sequence; Genes.  
PN JP 1989104167-A/2  
PD 21-APR-1989  
PF 01-AUG-1988 JP 1988192320  
PR 03-AUG-1987 GB 87 8718298, 26-OCT-1987 GB 87 8725052, PR  
13-NOV-1987 GB 87 8726683  
PI NIWA MINEO, SAITO YOSHIMASA, SASAKI HITOSHI, HAYASHI MASAKO,  
PI KOBAYASHI MASAKAZU  
PI NOTANI JOJI,  
PC C12N9/64, C12N1/20, C12N5/00, C12N15/00//A61K37/54, (C12N9/64, PC  
C12R1:19),  
PC (C12N9/64, C12R1:91);  
CC strandedness: Double;  
CC topology: Linear;  
CC hypothetical: No;  
CC \*source: clone=PTTKpadeltatrp;  
FH Key Location/Qualifiers  
FH FT  
CDS 1..1068  
/product='t-PA'.  
FEATURES  
source Location/Qualifiers  
1..1068  
/organism='synthetic construct'  
/mol\_type='genomic DNA'  
/db\_xref='taxon:32630'  
BASE COUNT 226 a 314 c 313 g 215 t  
ORIGIN  
Query Match 100.0%; Score 1065; DB 6; Length 1068;  
Best Local Similarity 100.0%; Pred. No. 4.7e-248;  
Matches 1065; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 TCTGAGGAAACAGTGACTGCTACTTTGGGAATGGGTGAGCTACCGTGGCAACGACAGC 60

Db 4 TCTGAGGAAACAGTGTCTTCTGGAAATGGGTGAGCTTACCGTGGACGACAGC 63  
Qy 61 CTCACCGAGTGGGTGCTCTCTGCTCCCGTGGAAATTCATGATCCTGATAGGCAAGTT 120  
Db 64 CTCACCGAGTGGGTGCTCTCTGCTCCCGTGGAAATTCATGATCCTGATAGGCAAGTT 123  
Qy 121 TACACAGACAGAACCCAGTGGCCAGGCACTGGGCTGGGCAACATAAATTAATCTGCGG 180  
Db 124 TACACAGACAGAACCCAGTGGCCAGGCACTGGGCTGGGCAACATAAATTAATCTGCGG 183  
Qy 181 AATCTGATGGGATGCCAAGCCCTGCTGACAGTGGTGGCAAGCCGAGCTGACGTTG 240  
Db 184 AATCTGATGGGATGCCAAGCCCTGCTGACAGTGGTGGCAAGCCGAGCTGACGTTG 243  
Qy 241 GAGTACTGTGATGTCCTCTCTGCTCCACCTGCGGCTGGACATGACAGCAGCTCAG 300  
Db 244 GAGTACTGTGATGTCCTCTCTGCTCCACCTGCGGCTGGACATGACAGCAGCTCAG 303  
Qy 301 TTTCGCATCAAGAGGAGCTCTTCGCGACATGCGCTCCACCCCTGCGGCTGACATC 360  
Db 304 TTTCGCATCAAGAGGAGCTCTTCGCGACATGCGCTCCACCCCTGCGGCTGACATC 363  
Qy 361 TTTCGCACAGCAGAGTGGCCCGGAGAGCGGTTCTCTGCGGGGGGATACTCATCAGC 420  
Db 364 TTTCGCACAGCAGAGTGGCCCGGAGAGCGGTTCTCTGCGGGGGGATACTCATCAGC 423  
Qy 421 TCCTGCTGATGTTCTCTGCGGCCACATGCTTCCAGGAGAGGTTTCCGCCCCACACCTG 480  
Db 424 TCCTGCTGATGTTCTCTGCGGCCACATGCTTCCAGGAGAGGTTTCCGCCCCACACCTG 483  
Qy 481 ACGTGATCTTTGGGAGAACATACCGGTGTTCCCTGCGGAGGAGGAGCAAGAAATTTGAA 540  
Db 484 ACGTGATCTTTGGGAGAACATACCGGTGTTCCCTGCGGAGGAGGAGCAAGAAATTTGAA 543  
Qy 541 GTCGAAATACATGTCATTAAGAAATTCATGATGACATTAAGCAATGACATTCGG 600  
Db 544 GTCGAAATACATGTCATTAAGAAATTCATGATGACATTAAGCAATGACATTCGG 603  
Qy 601 CTGCTGACGTGAATCGGATTCGTCGCTGTCGCGGAGGAGCAGCGTGTCCGCACT 660  
Db 604 CTGCTGACGTGAATCGGATTCGTCGCTGTCGCGGAGGAGCAGCGTGTCCGCACT 663  
Qy 661 GTGTGCTTCCCGGCGGACCTGACGTGCGGACCTGACGAGTGTGAGCTCTCCGCG 720  
Db 664 GTGTGCTTCCCGGCGGACCTGACGTGCGGACCTGACGAGTGTGAGCTCTCCGCG 723  
Qy 721 TACGCAAGCATGAGGCTTCTCTCTTCTATTCGAGCGGCTGAGAGGCTCATGTC 780  
Db 724 TACGCAAGCATGAGGCTTCTCTCTTCTATTCGAGCGGCTGAGAGGCTCATGTC 783  
Qy 781 AGACTGTACCATCCAGCGCTGCACATCAACATTTTACAGAAACAGTCAACGAC 840  
Db 784 AGACTGTACCATCCAGCGCTGCACATCAACATTTTACAGAAACAGTCAACGAC 843  
Qy 841 AACATGCTGTGTGGAGACACTCGGAGCGGCGGCGCCCGGCAAACTTGCACGAGCC 900  
Db 844 AACATGCTGTGTGGAGACACTCGGAGCGGCGGCGCCCGGCAAACTTGCACGAGCC 903  
Qy 901 TCCAGGCGGATTCGGGAGGCGGCTGTGTGTGTGAACGATGCGCGCATGATTTGGTG 960  
Db 904 TCCAGGCGGATTCGGGAGGCGGCTGTGTGTGTGAACGATGCGCGCATGATTTGGTG 963  
Qy 961 GGCATCATCAGTGGGCTGGTGTGACAGAGGATGTCGCGGCTGTACACAAAG 1020  
Db 964 GGCATCATCAGTGGGCTGGTGTGACAGAGGATGTCGCGGCTGTACACAAAG 1023  
Qy 1021 GTTACCAACTACCTAGATGATGATGACAAACATGCGACCGTGA 1065  
Db 1024 GTTACCAACTACCTAGATGATGATGACAAACATGCGACCGTGA 1068

RESULT 6

A27727  
LOCUS DNA sequence of coding region in pmTtk. linear PAT 09-JUL-2002  
ACCESSION A27727  
VERSION A27727.1 GI:21727237  
KEYWORDS  
SOURCE synthetic construct  
ORGANISM artificial construct  
REFERENCE 1 (bases 1 to 1170)  
AUTHORS Niwa, M., Saito, Y., Sasaki, H., Hayashi, M., Notani, J. and Kobayashi, M.  
TITLE New tissue plasminogen activator  
JOURNAL Patent: EP 0302456-A 51 08-FEB-1989;  
FUJISAWA PHARMACEUTICAL CO., LTD  
FEATURES  
Location/Qualifiers  
1..1170  
/organism="synthetic construct"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:32630"  
BASE COUNT 249 a 338 c 346 g 237 t  
ORIGIN

Query Match 100.0%; Score 1065; DB 6; Length 1170;  
Best Local Similarity 100.0%; Pred. No. 4.8e-248;  
Matches 1065; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 TCTCAGGGAACACGTGCTTCTTGGGAATGGTCCAGCTACCGTGGCAGCACAGC 60  
Db 106 TCTCAGGGAACACGTGCTTCTTGGGAATGGTCCAGCTACCGTGGCAGCACAGC 165  
Qy 61 CTCACCGAGTGGGTGCTCTCTGCTCCCGTGGAAATTCATGATCCTGATAGGCAAGTT 120  
Db 166 CTCACCGAGTGGGTGCTCTCTGCTCCCGTGGAAATTCATGATCCTGATAGGCAAGTT 225  
Qy 121 TACACAGCAGAAACCCAGTGGCCAGGCACTGGGCTGGGCAACATAAATTAATCTGCGG 180  
Db 226 TACACAGCAGAAACCCAGTGGCCAGGCACTGGGCTGGGCAACATAAATTAATCTGCGG 285  
Qy 181 AATCTCTATGGGATGCCAAGCCCTGCTGTCACCTGCTGAGAACCCGAGCTGACGTTG 240  
Db 286 AATCTCTATGGGATGCCAAGCCCTGCTGTCACCTGCTGAGAACCCGAGCTGACGTTG 345  
Qy 241 GAGTACTGTGATGTCCTCTCTGCTCCACCTGCGGCTGAGACAGTACAGCAGCTCAG 300  
Db 346 GAGTACTGTGATGTCCTCTCTGCTCCACCTGCGGCTGAGACAGTACAGCAGCTCAG 405  
Qy 301 TTTCGCATCAAGAGGAGGCTCTTCGCGGACATGCGCTCCACCCCTGGCAGGCTGCCATC 360  
Db 406 TTTCGCATCAAGAGGAGGCTCTTCGCGGACATGCGCTCCACCCCTGGCAGGCTGCCATC 465  
Qy 361 TTTCGCAAGCAGAGGAGTGGCCGAGAGCGGTTCTGTCGCGGGGCATACTCATCAGC 420  
Db 466 TTTCGCAAGCAGAGGAGTGGCCGAGAGCGGTTCTGTCGCGGGGCATACTCATCAGC 525  
Qy 421 TCCTGCTGGAATCTCTCTGCGGCCACTGCTTCAGGAGAGGTTTCGCGGCCACCACTG 480  
Db 526 TCCTGCTGGAATCTCTCTGCGGCCACTGCTTCAGGAGAGGTTTCGCGGCCACCACTG 585  
Qy 481 ACGTGATCTTTGGGCAAGACATACCGGCTGGTCCCTGGCGAGGAGGAGCAAAATTTGAA 540  
Db 586 ACGTGATCTTTGGGCAAGACATACCGGCTGGTCCCTGGCGAGGAGGAGCAAAATTTGAA 645  
Qy 541 GTCGAAATACATGTTGTCATAGGATTCGATGATGACACTTACGCAATGACATTCGCG 600  
Db 646 GTCGAAATACATGTTGTCATAGGATTCGATGATGACACTTACGCAATGACATTCGCG 705  
Qy 601 CTGCTGAGCTGAATCGGATTCGTCGCGCTGTGCCAGGAGAGCAGCGTGGTCCGCACT 660  
Db 706 CTGCTGAGCTGAATCGGATTCGTCGCGCTGTGCCAGGAGAGCAGCGTGGTCCGCACT 765  
Qy 661 GTGTGCTTCCCGGCGGACCTGCACTCCGAGCTGGAGCGGAGTGTGAGCTCTCCGCG 720





```
Db 1126 GTTACCAACTACCTAGATGGATTGCTGACAAATGCGACCGTGA 1170
|||||
RESULT 8
A27435 1314 bp DNA linear PAT 09-JUL-2002
DEFINITION DNA sequence of coding region in pTQKPADELtatr.
ACCESSION A27435
VERSION A27435.1 GI:21727230
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
REFERENCE 1 (bases 1 to 1314)
AUTHORS Niwa,M., Saito,Y., Sasaki,H., Hayashi,M., Notani,J. and Kobayashi,M.
TITLE New tissue plasminogen activator
JOURNAL Patent: EP 0302456-A 35 08-FEB-1989; FUJISAWA PHARMACEUTICAL CO., LTD
FEATURES
    source
        1. .1314
            /organism="synthetic construct"
            /mol_type="genomic DNA"
            /db_xref="caxon:32630"
BASE COUNT 286 a 386 c 393 g 249 t
ORIGIN
Query Match 100.0%; Score 1065; DB 6; Length 1314;
Best Local Similarity 100.0%; Pred. No. 4.8e-248; Mismatches 0; Indels 0; Gaps 0;
Matches 1065; Conservative 0;

Qy 1 TCTGAGGAAACAGTGAAGTCTGCTTTCGGGAATGGTCAAGCTACCGTGGCAGCAGCAGC 60
Db 250 TCTGAGGAAACAGTGAAGTCTGCTTTCGGGAATGGTCAAGCTACCGTGGCAGCAGCAGC 309
Qy 61 CTCACCGAGTGGGTGCTCTGCTCCCTCCGCTGGGAATTCATGATCCTCATAGGCAAGTT 120
Db 310 CTCACCGAGTGGGTGCTCTGCTCCCTCCGCTGGGAATTCATGATCCTCATAGGCAAGTT 369
Qy 121 TACACAGACAGAACCCAGTCCCGCAGCAGTGGCCCTGGCCGTCGACAAATATCTGCGG 180
Db 370 TACACAGACAGAACCCAGTCCCGCAGCAGTGGCCCTGGCCGTCGACAAATATCTGCGG 429
Qy 181 AATCTGATGGGATGCCAAGCCCTGCTGTCACAGCTGCTGAAGAACCGCAGCTGACGTGG 240
Db 430 AATCTGATGGGATGCCAAGCCCTGCTGTCACAGCTGCTGAAGAACCGCAGCTGACGTGG 489
Qy 241 GAGTACTGTGATGTGCTCTGCTCCACCTGCGGCTTGAGACAGTACAGCAGCAGCTCAG 300
Db 490 GAGTACTGTGATGTGCTCTGCTCCACCTGCGGCTTGAGACAGTACAGCAGCAGCTCAG 549
Qy 301 TTTGCGATCAAGAGGGGCTTTGCGCAGACATCGCTCCACCCCTGCGAGCTGCGCATC 360
Db 550 TTTGCGATCAAGAGGGGCTTTGCGCAGACATCGCTCCACCCCTGCGAGCTGCGCATC 609
Qy 361 TTTGCGAAGCAGAGAGTTCGCGGAGAGCGGTTCTGCTGCGGGGGCATCTCATCAGC 420
Db 610 TTTGCGAAGCAGAGAGTTCGCGGAGAGCGGTTCTGCTGCGGGGGCATCTCATCAGC 669
Qy 421 TCCTGCTGGATTCCTCTGCGGCCCACTGCTTCCAGGAGAGGTTTCCGCCCCCAACCTG 480
Db 670 TCCTGCTGGATTCCTCTGCGGCCCACTGCTTCCAGGAGAGGTTTCCGCCCCCAACCTG 729
Qy 481 ACGGTGATCTGGGAGACATACCGGTGCTCCCTGCGAGGAGGAGGACAGAAATTGAA 540
Db 730 ACGGTGATCTGGGAGACATACCGGTGCTCCCTGCGAGGAGGAGGACAGAAATTGAA 789
Qy 541 GTCGAAAAATACATTTGTCATAAGAAATTTCATGATGACACTTACGACAAATGACATTCGCG 600
Db 790 GTCGAAAAATACATTTGTCATAAGAAATTTCATGATGACACTTACGACAAATGACATTCGCG 849
Qy 601 CTGCTGCAGCTGAATCGGATTTCGTCCGCTGTGCCCCAGGAGCAGCGTGGTCCGCACT 660
|||||
Db 850 CTGCTGCAGCTGAATCGGATTTCGTCCGCTGTGCCAGGAGCAGCGTGGTCCGCACT 909
Qy 661 GTGTGCTTCCCGCGGACCTGAGCTGCGGAGCTGGACGAGTGTGAGCTCTCCGCGC 720
Db 910 GTGTGCTTCCCGCGGACCTGAGCTGCGGAGCTGGACGAGTGTGAGCTCTCCGCGC 969
Qy 721 TACGGCAAGCATGAGGCTTGTCTCTTTCTATTTCGAGCGGCTGAAGAGGCTCATGTC 780
Db 970 TACGGCAAGCATGAGGCTTGTCTCTTTCTATTTCGAGCGGCTGAAGAGGCTCATGTC 1029
Qy 781 AGACTGTACCATCCAGCGCTGCACATCAACATTTACTTTAAACAGACAGTACCGAC 840
Db 1030 AGACTGTACCATCCAGCGCTGCACATCAACATTTACTTTAAACAGACAGTACCGAC 1089
Qy 841 AACATGCTGTGTGCTGGAGACACTCGGAGCGGCGGCCCCAGGCAAACTTTCGACGAGCC 900
Db 1090 AACATGCTGTGTGCTGGAGACACTCGGAGCGGCGGCCCCAGGCAAACTTTCGACGAGCC 1149
Qy 901 TGCCAGGCGGATTCGGAGGCGCCCTGCTGTGCTGAACGATGCCCGCATGCTTTGGTG 960
Db 1150 TGCCAGGCGGATTCGGAGGCGCCCTGCTGTGCTGAACGATGCCCGCATGCTTTGGTG 1209
Qy 961 GGCATCATCAGCTGGGCGCTGGGCTGTGGACAGAGGATGTCCTCCGGTGTGTACACAAAG 1020
Db 1210 GGCATCATCAGCTGGGCGCTGGGCTGTGGACAGAGGATGTCCTCCGGTGTGTACACAAAG 1269
Qy 1021 GTTACCAACTACCTAGATGGATTTCGTGACAAATGCGACCGTGA 1065
Db 1270 GTTACCAACTACCTAGATGGATTTCGTGACAAATGCGACCGTGA 1314
|||||
RESULT 9
E01937 1314 bp DNA linear PAT 29-SEP-1997
LOCUS Synthetic DNA encoding new t-PA tissue plasminogen activator.
DEFINITION E01937
ACCESSION E01937
VERSION E01937.1 GI:2170185
KEYWORDS JP 1989104167-A/4.
SOURCE synthetic construct
ORGANISM synthetic construct
REFERENCE 1 (bases 1 to 1314)
AUTHORS Niwa,M., Saito,Y., Sasaki,H., Hayashi,M., Notani,J. and Kobayashi,M.
TITLE NOVEL TISSUE PLASMINOGEN ACTIVATION FACTOR
JOURNAL Patent: JP 1989104167-A 4 21-APR-1989; FUJISAWA PHARMACEUT CO LTD
COMMENT OS Artificial gene
OC Artificial sequence; Genes.
PN JP 1989104167-A/4
PD 21-APR-1989
PF 01-AUG-1988 JP 1988192320
PR 03-AUG-1987 GB 87 8718298, 26-OCT-1987 GB 87 8725052, PR 13-NOV-1987 GB 87 8726683
PI NIWA MINEO, SAITO YOSHIMASA, SASAKI HITOSHI, HAYASHI MASAKO, PI NOTANI JOJI,
PI KOBAYASHI MASAKAZU
PC C12N9/64, C12N1/20, C12N5/00, C12N15/00, C12N13/754, (C12N9/64, PC C12R1/19),
PC (C12N9/64, C12R1/91);
CC strandedness: Double;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
CC *source: clone=pTQKPADELtatr;
FH Key Location/Qualifiers
FT CDS 1. .1314
FT /product='t-PA'.
FT Location/Qualifiers
1. .1314
/organism='synthetic construct'
```

```
/mol_type="genomic DNA"
/db_xref="taxon:32630"
BASE COUNT      286 a      386 c      393 g      249 t
ORIGIN
Query Match      100.0%; Score 1065; DB 6; Length 1314;
Best Local Similarity 100.0%; Pred. No. 4.8e-248;
Matches 1065; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 TCTGAGGAAACAGTACTGCTTCTTGGGAATGGTTCAGCTTACCGTGGCAGCAGC 60
Db 250 TCTGAGGAAACAGTACTGCTTCTTGGGAATGGTTCAGCTTACCGTGGCAGCAGC 309
Qy 61 CTCACCGAGTGGGCTCTCTGCTCCCTCCGCTGGGAATTCATGATCCTCATAGGCAAGTT 120
Db 310 CTCACCGAGTGGGCTCTCTGCTCCCTCCGCTGGGAATTCATGATCCTCATAGGCAAGTT 369
Qy 121 TACACAGACAGAACCCAGTCCAGGCACTGGGCTGGGCAACATAATTAATCTGCGG 180
Db 370 TACACAGACAGAACCCAGTCCAGGCACTGGGCTGGGCAACATAATTAATCTGCGG 429
Qy 181 AATCCTGATGGGATGCCAAGCCCTGCTGCACGCTGCTGAGAACCCGAGCTGACGTGG 240
Db 430 AATCCTGATGGGATGCCAAGCCCTGCTGCACGCTGCTGAGAACCCGAGCTGACGTGG 489
Qy 241 GAGTACTGTGATGTGCTCTCTGCTCCACCTGCGGCTGAGACAGTACAGCAGCCTCAG 300
Db 490 GAGTACTGTGATGTGCTCTCTGCTCCACCTGCGGCTGAGACAGTACAGCAGCCTCAG 549
Qy 301 TTTGCGATCAAGAGGAGCTCTTTCGCGACATGCGCTCCACCCCTGCGAGGCTGCCATC 360
Db 550 TTTGCGATCAAGAGGAGCTCTTTCGCGACATGCGCTCCACCCCTGCGAGGCTGCCATC 609
Qy 361 TTTGCGACGACAGAGCTGCGCGAGAGCGGTTCTGTCGGGGGATACTCATCAGC 420
Db 610 TTTGCGACGACAGAGCTGCGCGAGAGCGGTTCTGTCGGGGGATACTCATCAGC 669
Qy 421 TCCTGCTGATTTCTCTGCGCGCCACCTGCTTCCAGAGAGGTTTCCGCCCCACCACTG 480
Db 670 TCCTGCTGATTTCTCTGCGCGCCACCTGCTTCCAGAGAGGTTTCCGCCCCACCACTG 729
Qy 481 ACGGTGATCTTGGGAGAACATACCGGGTGTCTTCCAGAGAGGAGAGCAAAATTTGAA 540
Db 730 ACGGTGATCTTGGGAGAACATACCGGGTGTCTTCCAGAGAGGAGAGCAAAATTTGAA 789
Qy 541 GTCGAAATATCATTTGTCATAGGAATTCGATGATGACATTACGACAAATGACATTGCG 600
Db 790 GTCGAAATATCATTTGTCATAGGAATTCGATGATGACATTACGACAAATGACATTGCG 849
Qy 601 CTGCTGACAGTGAATTCGGATTCGCTGCTGCTGCGGAGAGCAGCGTGGTCCGCACT 660
Db 850 CTGCTGACAGTGAATTCGGATTCGCTGCTGCGGAGAGCAGCGTGGTCCGCACT 909
Qy 661 GTGTGCTTCCCGCGGAGACTGAGCTGCGGACTGGAACGAGTGTGAGCTCTCCGCG 720
Db 910 GTGTGCTTCCCGCGGAGACTGAGCTGCGGACTGGAACGAGTGTGAGCTCTCCGCG 969
Qy 721 TACGCGACGATGAGGCTTGTCTCTTCTTATTCGAGCGGCTGAGAGGCTCATGTC 780
Db 970 TACGCGACGATGAGGCTTGTCTCTTCTTATTCGAGCGGCTGAGAGGCTCATGTC 1029
Qy 781 AGACTGTACCCATCCAGCGCTGACATCAACAATTTTACTTAAACAGAACAGTCAACGAC 840
Db 1030 AGACTGTACCCATCCAGCGCTGACATCAACAATTTTACTTAAACAGAACAGTCAACGAC 1089
Qy 841 AACATGCTGTGTGTGGAGACATCCGAGCGCGCGGCCCGGCAAACTTGACAGCGCC 900
Db 1090 AACATGCTGTGTGTGGAGACATCCGAGCGCGCGGCCCGGCAAACTTGACAGCGCC 1149
Qy 901 TCCAGGCGGATTCGGGAGGCGCCCTGCTGTGTGTAACGATGCGCGCATGACTTTGGTG 960
Db 1150 TCCAGGCGGATTCGGGAGGCGCCCTGCTGTGTGTAACGATGCGCGCATGACTTTGGTG 1209
```

---

```
Qy 961 GGCATCATCAGCTGGGCTGGCTGTGGACAGAGGATGTCCGGGTGTGTACACAAAG 1020
Db 1210 GGCATCATCAGCTGGGCTGGCTGTGGACAGAGGATGTCCGGGTGTGTACACAAAG 1269
Qy 1021 GTTACCAACTACCTAGACTGGATTCTGTGACAAACATGTCGACCGTGA 1065
Db 1270 GTTACCAACTACCTAGACTGGATTCTGTGACAAACATGTCGACCGTGA 1314

RESULT 10
A27725
LOCUS      A27725      1419 bp      DNA      linear      PAT 09-JUL-2002
DEFINITION DNA sequence of coding region in pmfQk112.
ACCESSION  A27725
VERSION     A27725.1 GI:21727236
KEYWORDS    .
SOURCE      synthetic construct
ORGANISM    synthetic construct
            artificial sequences.
REFERENCE   1 (bases 1 to 1419)
            Niwa,M., Saito,Y., Sasaki,H., Hayaashi,M., Notani,J. and
            Kobayashi,M.
            New tissue plasminogen activator
            Patent: BP 0302456-A 49 08-FEB-1989;
            FUJISAWA PHARMACEUTICAL CO., LTD
FEATURES
            Location/Qualifiers
            1..1419
            /organism="synthetic construct"
            /mol_type="genomic DNA"
            /db_xref="taxon:32630"
BASE COUNT  309 a      413 c      426 g      271 t
ORIGIN
Query Match      100.0%; Score 1065; DB 6; Length 1419;
Best Local Similarity 100.0%; Pred. No. 4.8e-248;
Matches 1065; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 TCTGAGGAAACAGTACTGCTTCTTGGGAATGGGTCAGCTTACCGTGGCAGCAGC 60
Db 355 TCTGAGGAAACAGTACTGCTTCTTGGGAATGGGTCAGCTTACCGTGGCAGCAGC 414
Qy 61 CTCACCGAGTGGGCTCTCTGCTCCCTCCGCTGGGAATTCATGATCTCTGATAGCAAGTT 120
Db 415 CTCACCGAGTGGGCTCTCTGCTCCCTCCGCTGGGAATTCATGATCTCTGATAGCAAGTT 474
Qy 121 TACACAGACAGAACCCAGTGGCAGGCACTGGGCTGGGCAACATAATTAATCTGCGG 180
Db 475 TACACAGACAGAACCCAGTGGCAGGCACTGGGCTGGGCAACATAATTAATCTGCGG 534
Qy 181 AATCCTGATGGGATGCCAAGCCCTGCTGCGACGCTGCTGAGAACCCGAGCTGACGTGG 240
Db 535 AATCCTGATGGGATGCCAAGCCCTGCTGCGACGCTGCTGAGAACCCGAGCTGACGTGG 594
Qy 241 GAGTACTGTGATGTGCTCTCTGCTCCACCTGCGGCTGAGACAGTACAGCAGCCTCAG 300
Db 595 GAGTACTGTGATGTGCTCTCTGCTCCACCTGCGGCTGAGACAGTACAGCAGCCTCAG 654
Qy 301 TTTGCGATCAAGAGGAGGCTTTCGCGGACATGCGCTCCACCCCTGGCAGGCTGCGCATC 360
Db 655 TTTGCGATCAAGAGGAGGCTTTCGCGGACATGCGCTCCACCCCTGGCAGGCTGCGCATC 714
Qy 361 TTTGCGAAGCAGAGGAGGCTGCGCGGAGAGCGGTTCTGTGCGGGGACATCTCATCAGC 420
Db 715 TTTGCGAAGCAGAGGAGGCTGCGCGGAGAGCGGTTCTGTGCGGGGACATCTCATCAGC 774
Qy 421 TCCTGCTGGAATTTCTCTGCGCGCCACTGCTTCCAGGAGAGGTTTCCGCCCCACCACTG 480
Db 775 TCCTGCTGGAATTTCTCTGCGCGCCACTGCTTCCAGGAGAGGTTTCCGCCCCACCACTG 834
Qy 481 ACGGTGATTTGGGAGAACATATCCGGGTGGTTCCTGCGAGGAGGAGCAAAATTTGAA 540
Db 835 ACGGTGATTTGGGAGAACATATCCGGGTGGTTCCTTGGCGAGGAGGAGCAAAATTTGAA 894
```

	Query Match	100.0%;	Score 1065;	DB 6;	Length 1419;
	Best Local Similarity	100.0%;	Pred. No. 4.8e-248;		
	Matches 1065;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	TCTGAGGGAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTGCGACGCAAGC	60		
Db	355				
		TCTGAGGGAACAGTGACTGCTACTTTGGGAATGGGTCAGCCTACCGTGCGACGCAAGC	414		
Qy	61	CTCACCGAGTCGGGTGCCTCTGCTCTCCCGTGGAAATCCATGATCTCTGATAGCGCAAGTT	120		
Db	415				
		CTCACCGAGTCGGGTGCCTCTGCTCTCCCGTGGAAATCCATGATCTCTGATAGCGCAAGTT	474		
Qy	121	TACACGCA CAGAACCCCGATGCCCGCCAGGCACTGGCGCTGGGGCAAAACAATAATTACTGCCGG	180		
Db	475				
		TACACGCA CAGAACCCCGATGCCCGCCAGGCACTGGCGCTGGGGCAAAACAATAATTACTGCCGG	534		

RESULT 12  
E01944  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE

AUTHORS Niwa,M., Saito,Y., Sasaki,H., Hayashi,M., Notani,J. and Kobayashi,M.  
 TITLE NOVEL TISSUE PLASMINOGEN ACTIVATION FACTOR  
 JOURNAL Patent: JP 1989104167-A 11 21-APR-1989;  
 FUJISAWA PHARMACEUT CO LTD  
 COMMENT OS Artificial gene  
 OC Artificial sequence; Genes.  
 PD 21-APR-1989  
 PN JP 1989104167-A/11  
 PF 01-AUG-1988 JP 1988192320  
 PR 03-AUG-1987 GB 87 8718298, 26-OCT-1987 GB 87 8725052, PR 13-NOV-1987 GB 87 8726683  
 PI NIWA MINO, SAITO YOSHIMASA, SASAKI HITOSHI, HAYASHI MASAKO, NOTANI JOJI.  
 PI KOBAYASHI MASAKAZU  
 PC C12N9/64, C12N1/20, C12N5/00, C12N15/00//A61K37/54, (C12N9/64, PC C12R1:19),  
 PC (C12N9/64, C12R1:91);  
 CC strandedness: Double;  
 CC topology: Linear;  
 CC hypothetical: No;  
 CC anti-sense: No;  
 CC \*source: clone=pmTqk112;  
 FH Key Location/Qualifiers  
 FT CDS 1..1419 /product=new t-PA'.  
 FT Location/Qualifiers  
 1..1419  
 /organism="synthetic construct"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:32630"  
 BASE COUNT 309 a 413 c 426 g 271 t  
 ORIGIN

Query Match 100.0%; Score 1065; DB 6; Length 1419;  
 Best Local Similarity 100.0%; Pred. No. 4.8e-248;  
 Matches 1065; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCTGAGGAAACAGTACTGCTCTTGGGAATGGGTGAGCTACCGTGGGACGACAGC 60  
 Db 355 TCTGAGGAAACAGTACTGCTCTTGGGAATGGGTGAGCTACCGTGGGACGACAGC 414  
 Qy 61 CTCACCGAGTGGGTGCTCTCTGCTCCGCTGGAATTCATGATCCTCATAGGCAAGTT 120  
 Db 415 CTCACCGAGTGGGTGCTCTCTGCTCCGCTGGAATTCATGATCCTCATAGGCAAGTT 474  
 Qy 121 TACACGACAGAACCCAGTCCGAGGACCTGGGCTGGGCAACACATAATTAATCTGCCG 180  
 Db 475 TACACGACAGAACCCAGTCCGAGGACCTGGGCTGGGCAACACATAATTAATCTGCCG 534  
 Qy 181 AATCTGTAGTGGGATGCCAGCCCTGGTCCGCTGGAAGTGGTGAAGACCGGAGCTGACGTGG 240  
 Db 535 AATCTGTAGTGGGATGCCAGCCCTGGTCCGCTGGAAGTGGTGAAGACCGGAGCTGACGTGG 594  
 Qy 241 GAGTACTGTGATGTCCTCTCTGCTCCACCTGCGGCTGAGACAGTACAGCCAGCTCAG 300  
 Db 595 GAGTACTGTGATGTCCTCTCTGCTCCACCTGCGGCTGAGACAGTACAGCCAGCTCAG 654  
 Qy 301 TTTCGCATCAAGAGGAGGCTCTTCGCGACATCGCTCCACCCCTGGCAGGCTGCCATC 360  
 Db 655 TTTCGCATCAAGAGGAGGCTCTTCGCGACATCGCTCCACCCCTGGCAGGCTGCCATC 714  
 Qy 361 TTTCGCAAGCAGAGAGGTGCGCCGAGAGCGGTTCTGTGGGGGGGATATCATCAGC 420  
 Db 715 TTTCGCAAGCAGAGAGGTGCGCCGAGAGCGGTTCTGTGGGGGGGATATCATCAGC 774  
 Qy 421 TCCTGCTGATTCCTCTCTGCGGCCCACTGCTTCCAGGAGAGGTTTCCGCCCCACACCTG 480  
 Db 775 TCCTGCTGATTCCTCTCTGCGGCCCACTGCTTCCAGGAGAGGTTTCCGCCCCACACCTG 834  
 Qy 481 ACGGTGATCTGGGACAGACATACCGGCTGCTCTGCGGAGGAGGACAGAAATTTGAA 540

Db 835 ACGGTGATCTTGGGACAGACATACCGGCTGCTCTGCGAGGAGGACAGAAATTTGAA 894  
 Qy 541 GTCGAAAATACATTTGTCATTAAGGAATTCGATGATGACATTTACGCAATGACATTCGC 600  
 Db 895 GTCGAAAATACATTTGTCATTAAGGAATTCGATGATGACATTTACGCAATGACATTCGC 954  
 Qy 601 CTGCTGCAGCTGAAATCGGATTCGTCCTGCTGTCCTCCAGAGAGCAGCGTGGTCCGCACT 660  
 Db 955 CTGCTGCAGCTGAAATCGGATTCGTCCTGCTGTCCTCCAGAGAGCAGCGTGGTCCGCACT 1014  
 Qy 661 GTGTCCTTCCCCCGGCGGACCTGACGTGCGGACCTGGACGGAGTGTGAGCTTCCGGC 720  
 Db 1015 GTGTCCTTCCCCCGGCGGACCTGACGTGCGGACCTGGACGGAGTGTGAGCTTCCGGC 1074  
 Qy 721 TACGGCAAGCATGAGGCTTGTCTCTCTTCTATTCGAGCGGCTGAGAGGCTCATGTC 780  
 Db 1075 TACGGCAAGCATGAGGCTTGTCTCTCTTCTATTCGAGCGGCTGAGAGGCTCATGTC 1134  
 Qy 781 AGACTGTACCCATCCAGCCGCTGCACATCAACATTTACTTTACAGAACAGTCAACCGAC 840  
 Db 1135 AGACTGTACCCATCCAGCCGCTGCACATCAACATTTACTTTACAGAACAGTCAACCGAC 1194  
 Qy 841 AACATGCTGTGCTGGAGACATTCGAGCGGCGGCCCCAGGCAAACTTGCACAGCGCC 900  
 Db 1195 AACATGCTGTGCTGGAGACATTCGAGCGGCGGCCCCAGGCAAACTTGCACAGCGCC 1254  
 Qy 901 TGCAGGCGCATTCGGAGGCGCCCTGCTGCTGCTGAACGATGGCCGATGACTTTGGTG 960  
 Db 1255 TGCAGGCGCATTCGGAGGCGCCCTGCTGCTGCTGAACGATGGCCGATGACTTTGGTG 1314  
 Qy 961 GGCATCATCAGCTGGGCGCTGGGCTGGGACAGAGGATGTCCCGGTGTGTACACAAAG 1020  
 Db 1315 GGCATCATCAGCTGGGCGCTGGGCTGGGACAGAGGATGTCCCGGTGTGTACACAAAG 1374  
 Qy 1021 GTTACCAACTACCTAGACTGGATTCGTGACAAATCGACCGCTGA 1065  
 Db 1375 GTTACCAACTACCTAGACTGGATTCGTGACAAATCGACCGCTGA 1419

RESULT 13  
 LOCUS 106609 1689 bp DNA linear PAT 02-DEC-1994  
 DEFINITION Sequence 44 from Patent EP 0293934.  
 ACCESSION I06609  
 VERSION I06609.1 GI:590639  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unknown.  
 REFERENCES  
 1 (bases 1 to 1689)  
 Mulvihill, E.R., Nexo, B.A., Yoshitake, S., Ikeda, Y., Suzuki, S., Hashimoto, A. and Yuzuriha, T.  
 Mutant t-PA with kringle replacement  
 Patent: EP 0293934-A1 44 07-DEC-1988;  
 JOURNAL  
 FEATURES  
 source 1..1689  
 /organism="unknown"  
 BASE COUNT 377 a 483 c 505 g 324 t  
 ORIGIN

Query Match 100.0%; Score 1065; DB 6; Length 1689;  
 Best Local Similarity 100.0%; Pred. No. 4.8e-248;  
 Matches 1065; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCTGAGGAAACAGTACTGCTCTTGGGAATGGGTGAGCTACCGTGGGACGACAGC 60  
 Db 625 TCTGAGGAAACAGTACTGCTCTTGGGAATGGGTGAGCTACCGTGGGACGACAGC 684  
 Qy 61 CTCACCGAGTGGGTGCTCTCTGCTCCGCTGGAATTCATGATCCTCATAGGCAAGTT 120  
 Db 685 CTCACCGAGTGGGTGCTCTCTGCTCCGCTGGAATTCATGATCCTCATAGGCAAGTT 744  
 Qy 121 TACACGACAGAACCCAGTCCGAGGACCTGGGCTGGGCAACACATAATTAATCTGCCG 180

Db 745 TACACAGACAGAAACCCAGTGCAGGACCTGGCCCTGGGCAACATAATTTACTGCCGG 804  
Qy 181 AATCTGATGGGATGCCAAGCCCTGTGTGCCACAGTGTGAAGAACCGCAGGCTGACGTGG 240  
Db 805 AATCTGATGGGATGCCAAGCCCTGTGTGCCACAGTGTGAAGAACCGCAGGCTGACGTGG 864  
Qy 241 GAGTACTGTGATGTGCTCTCTGTCTCCACCTGTGCGGCTGTGAGACAGTACAGCGAGCTCAG 300  
Db 865 GAGTACTGTGATGTGCTCTCTGTCTCCACCTGTGCGGCTGTGAGACAGTACAGCGAGCTCAG 924  
Qy 301 TTTGSCATCAAGAGAGGCTCTTCCGCGACATCGCTCCACCCCTGCGAGGCTGCCATC 360  
Db 925 TTTGSCATCAAGAGAGGCTCTTCCGCGACATCGCTCCACCCCTGCGAGGCTGCCATC 984  
Qy 361 TTTGSCAAGCAGAGAGGCTCTTCCGCGACATCGCTCCACCCCTGCGAGGCTGCCATC 420  
Db 985 TTTGSCAAGCAGAGAGGCTCTTCCGCGACATCGCTCCACCCCTGCGAGGCTGCCATC 1044  
Qy 421 TCCTGCTGGATCTCTCTGCGGCTCTCTGCTTCCAGGAGAGGTTTCCGCGGCTGCCATC 480  
Db 1045 TCCTGCTGGATCTCTCTGCGGCTCTCTGCTTCCAGGAGAGGTTTCCGCGGCTGCCATC 1104  
Qy 481 ACAGTGTATCTGGGAGAGGCTCTTCCGCGACATCGCTCCACCCCTGCGAGGCTGCCATC 540  
Db 1105 ACAGTGTATCTGGGAGAGGCTCTTCCGCGACATCGCTCCACCCCTGCGAGGCTGCCATC 1164  
Qy 541 CTCGAAATATACATTTGTCATTAAGAAATTCGATGATGACATTAACGAAATGACATTTG 600  
Db 1165 CTCGAAATATACATTTGTCATTAAGAAATTCGATGATGACATTAACGAAATGACATTTG 1224  
Qy 601 CTGCTGAGCTGAATTCGATGATGATGATGATGATGATGATGATGATGATGATGATG 660  
Db 1225 CTGCTGAGCTGAATTCGATGATGATGATGATGATGATGATGATGATGATGATGATG 1284  
Qy 661 GTGTGCTTCCCGCGGAGCTGCGAGCTGCGGAGCTGCGAGCTGCGAGCTGCGAGCTGCG 720  
Db 1285 GTGTGCTTCCCGCGGAGCTGCGAGCTGCGGAGCTGCGAGCTGCGAGCTGCGAGCTGCG 1344  
Qy 721 TACGCAAGCATGAGGCTTGTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 780  
Db 1345 TACGCAAGCATGAGGCTTGTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1404  
Qy 781 AGACTGTACCATCCAGCCGCTGCACATCAACAATTTACTTAACAGAACAGTCAACGAC 840  
Db 1405 AGACTGTACCATCCAGCCGCTGCACATCAACAATTTACTTAACAGAACAGTCAACGAC 1464  
Qy 841 AACATGCTGT 900  
Db 1465 AACATGCTGT 1524  
Qy 901 TGCCAGGCGGATTCGGAGGCGCCCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 960  
Db 1525 TGCCAGGCGGATTCGGAGGCGCCCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1584  
Qy 961 GGCATCATCAGCTGGGCTGTGGAGTGTGGAGAGGATGTCCCGGCTGTGTACACAAAG 1020  
Db 1585 GGCATCATCAGCTGGGCTGTGGAGTGTGGAGAGGATGTCCCGGCTGTGTACACAAAG 1644  
Qy 1021 GTTACCACTACCTAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1065  
Db 1645 GTTACCACTACCTAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1689

RESULT 14  
106614 I06614 Sequence 46 from Patent EP 0293934.  
LOCUS I06614  
DEFINITION I06614  
ACCESSION I06614  
VERSION I06614.1  
GI:590640  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.

Unclassified.  
1 (bases 1 to 1689)  
REFERENCE Mulvihill, E.R., Nexo, B.A., Yoshitake, S., Ikeda, Y., Suzuki, S.,  
AUTHORS Hashimoto, A. and Yuzuriha, T.  
TITLE Mutant t-Pa with kringle replacement  
JOURNAL Patent: EP 0293934-A1 46 07-DEC-1988;  
FEATURES Location/Qualifiers  
source 1.1689  
BASE COUNT 377 a 484 c 504 g 324 t  
ORIGIN  
Query Match 100.0%; Score 1065; DB 6; Length 1689;  
Best Local Similarity 100.0%; Pred. No. 4.8e-248;  
Matches 1065; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 TCTGAGGGAACAGTGAAGTCTTCTTGGGAATGGGTGAGCTTACCTGGCAGCCACAGC 60  
Db 625 TCTGAGGGAACAGTGAAGTCTTCTTGGGAATGGGTGAGCTTACCTGGCAGCCACAGC 684  
Qy 61 CTCAACGAGTGGGTGCTCTCTGCTCCCGTGGAAATTCATGATCTCTGATAGGCAAGTT 120  
Db 685 CTCAACGAGTGGGTGCTCTCTGCTCCCGTGGAAATTCATGATCTCTGATAGGCAAGTT 744  
Qy 121 TACACAGACAGAAACCCAGTGCAGGCACTGGGCTTGGGCAACATAATTTACTGCGG 180  
Db 745 TACACAGACAGAAACCCAGTGCAGGCACTGGGCTTGGGCAACATAATTTACTGCGG 804  
Qy 181 AATCTGATGGGATGCCAAGCCCTGTGTGCCAGTGTGAGAACCCGAGCTGACGTGG 240  
Db 805 AATCTGATGGGATGCCAAGCCCTGTGTGCCAGTGTGAGAACCCGAGCTGACGTGG 864  
Qy 241 GAGTACTGTGATGTGCTCTCTGCTCCACCTGCGGCTGAGACAGTACAGCCACCTCAG 300  
Db 865 GAGTACTGTGATGTGCTCTCTGCTCCACCTGCGGCTGAGACAGTACAGCCACCTCAG 924  
Qy 301 TTTGSCATCAAGAGAGGCTCTTCCGCGACATCGCTCCACCCCTGCGAGGCTGCCATC 360  
Db 925 TTTGSCATCAAGAGAGGCTCTTCCGCGACATCGCTCCACCCCTGCGAGGCTGCCATC 984  
Qy 361 TTTGSCAAGCAGAGAGGCTCTTCCGCGACATCGCTCCACCCCTGCGAGGCTGCCATC 420  
Db 985 TTTGSCAAGCAGAGAGGCTCTTCCGCGACATCGCTCCACCCCTGCGAGGCTGCCATC 1044  
Qy 421 TCCTGCTGGATTTCTCTGCGGCTCTCTGCTTCCAGGAGAGGTTTCCGCGGCTGCCATC 480  
Db 1045 TCCTGCTGGATTTCTCTGCGGCTCTCTGCTTCCAGGAGAGGTTTCCGCGGCTGCCATC 1104  
Qy 481 ACAGTGTATCTGGGAGAGGCTCTTCCGCGACATCGCTCCAGGAGGAGGAGAAATTTGAA 540  
Db 1105 ACAGTGTATCTGGGAGAGGCTCTTCCGCGACATCGCTCCAGGAGGAGGAGAAATTTGAA 1164  
Qy 541 GTCGAAATATACATTTGTCATTAAGAAATTCGATGATGACATTAACGAAATGACATTTG 600  
Db 1165 GTCGAAATATACATTTGTCATTAAGAAATTCGATGATGACATTAACGAAATGACATTTG 1224  
Qy 601 CTGCTGAGCTGAATTCGATGATGATGATGATGATGATGATGATGATGATGATGATG 660  
Db 1225 CTGCTGAGCTGAATTCGATGATGATGATGATGATGATGATGATGATGATGATGATG 1284  
Qy 661 GTGTGCTTCCCGCGGAGCTGCGAGCTGCGGAGCTGCGAGCTGCGAGCTGCGAGCTGCG 720  
Db 1285 GTGTGCTTCCCGCGGAGCTGCGAGCTGCGGAGCTGCGAGCTGCGAGCTGCGAGCTGCG 1344  
Qy 721 TACGCAAGCATGAGGCTTGTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 780  
Db 1345 TACGCAAGCATGAGGCTTGTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1404  
Qy 781 AGACTGTACCATCCAGCCGCTGCACATCAACAATTTACTTAACAGAACAGTCAACGAC 840  
Db 1405 AGACTGTACCATCCAGCCGCTGCACATCAACAATTTACTTAACAGAACAGTCAACGAC 1464  
Qy 841 AACATGCTGT 900

Db 1465 AACATGCTGTGTGTGAGACACTCGGAGCGCGCGCCAGGCAAACTTGACGACGCC 1524  
Qy 901 TGCCAGGCGGATTCGGGAGGCCCTCGTGTGTCTGAAAGATGCGCGCATGACTTTGGTG 960  
Db 1525 TGCCAGGCGGATTCGGGAGGCCCTCGTGTGTCTGAAAGATGCGCGCATGACTTTGGTG 1584  
Qy 961 GGCATCATCAGCTGGGGCTCGGGCTGTGACAGAGAGATGTCCCGGGTGTGTACACAAAG 1020  
Db 1585 GGCATCATCAGCTGGGGCTCGGGCTGTGACAGAGAGATGTCCCGGGTGTGTACACAAAG 1644  
Qy 1021 GTTACCAACTACTAGATGGAATTCGTGACAACTGCAACCGTGA 1065  
Db 1645 GTTACCAACTACTAGATGGAATTCGTGACAACTGCAACCGTGA 1689

RESULT 15  
I08789  
LOCUS I08789 1689 bp DNA linear PAT 02-DEC-1994  
DEFINITION Sequence 3 from Patent WO 8804690.  
ACCESSION I08789  
VERSION I08789.1 GI:588494  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 1689)  
Unclassified.  
AUTHORS Bollen,A.J., Gheysen,D., Jacobs,P., Pierard,L. and Collen,D.J.  
JOURNAL Patent: WO 8804690-A 3 30-JUN-1988;  
FEATURES  
Location/Qualifiers  
source 1..1689  
/organism="unknown"  
BASE COUNT 379 a 480 c 505 g 325 t  
ORIGIN

Query Match 100.0%; Score 1065; DB 6; Length 1689;  
Best Local Similarity 100.0%; Pred. No. 4.8e-248;  
Matches 1065; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCTGAGGAAACAGTACTGTCTACTTTGGGATGGGTGAGCTACCGTGGGACGACAGC 60  
Db 625 TCTGAGGAAACAGTACTGTCTACTTTGGGATGGGTGAGCTACCGTGGGACGACAGC 684  
Qy 61 CTCACCGAGTCGGGTGCTCTGCTCCCGTGGGAATCCATGATCCTGATAGGCAAGTT 120  
Db 685 CTCACCGAGTCGGGTGCTCTGCTCCCGTGGGAATCCATGATCCTGATAGGCAAGTT 744  
Qy 121 TACACAGCACAGAACCCAGTGCCGACGCACTGGGCTGGGCAAAACATAATTAATCTGCCG 180  
Db 745 TACACAGCACAGAACCCAGTGCCGACGCACTGGGCTGGGCAAAACATAATTAATCTGCCG 804  
Qy 181 AATCCTGATGGGATGCCAAGCCCTGGTGCCACGTCGTGAAGAACCGAGGCTGACGTGG 240  
Db 805 AATCCTGATGGGATGCCAAGCCCTGGTGCCACGTCGTGAAGAACCGAGGCTGACGTGG 864  
Qy 241 GAGTACTGTGATGTCCTCTGCTCCACCTGCGGCTGAGACAGTACAGCAGCCTCAG 300  
Db 865 GAGTACTGTGATGTCCTCTGCTCCACCTGCGGCTGAGACAGTACAGCAGCCTCAG 924  
Qy 301 TTTGCGCATCAAGAGGAGGCTCTTGCCGACATGCGCTCCCAACCCCTGGCAGGCTGCCATC 360  
Db 925 TTTGCGCATCAAGAGGAGGCTCTTGCCGACATGCGCTCCCAACCCCTGGCAGGCTGCCATC 984  
Qy 361 TTTGCGCAGCAGCAGAGTTCGCGGAGAGCGGTTCCTGTGCGGGGATCTATCATCAGC 420  
Db 985 TTTGCGCAGCAGCAGAGTTCGCGGAGAGCGGTTCCTGTGCGGGGATCTATCATCAGC 1044  
Qy 421 TCCTGCTGATTCCTCTGCGCGCCACCTGCTTCAGAGAGAGGTTTCGCGCCCAACACCTG 480  
Db 1045 TCCTGCTGATTCCTCTGCGCGCCACCTGCTTCAGAGAGAGGTTTCGCGCCCAACACCTG 1104  
Qy 481 ACGTGTATCTTGGGACAGAACATACCGGTGTGTCCTGCGGAGGAGGACAGAAATTTGAA 540

Search completed: August 19, 2003, 04:46:33  
Job time : 4473.78 secs

Db 1105 ACGTGTATCTTTGGCAGACATACCGGTGTCTCCGCGAGGAGGACGAAATTTGAA 1164  
Qy 541 GTCGAAAAATACATTTGTCCATAAGGAATTCGATGATGACACTTACGACAATGACATTCGCG 600  
Db 1165 GTCGAAAAATACATTTGTCCATAAGGAATTCGATGATGACACTTACGACAATGACATTCGCG 1224  
Qy 601 CTGCTGCAGCTGAATTCGGATTCGTCCCGTGTGTCGCCAGGAGAGCAGCGTGGTCCGCACT 660  
Db 1225 CTGCTGCAGCTGAATTCGGATTCGTCCCGTGTGTCGCCAGGAGAGCAGCGTGGTCCGCACT 1284  
Qy 661 GTGTGCTTCCCGCGGACCTGCGAGCTGCGGACCTGGAACGAGTGTGAGCTCTCCGGC 720  
Db 1285 GTGTGCTTCCCGCGGACCTGCGAGCTGCGGACCTGGAACGAGTGTGAGCTCTCCGGC 1344  
Qy 721 TAGGCAAGCATGAGGCTTTGTCTCTTCTATTTCGGAGCGGCTGAAGGAGGCTCATGTC 780  
Db 1345 TAGGCAAGCATGAGGCTTTGTCTCTTCTATTTCGGAGCGGCTGAAGGAGGCTCATGTC 1404  
Qy 781 AGACTGTACCCATCCAGCCGCTGCACATCACACATTTACTTTAACAGAACAGTCACCGAC 840  
Db 1405 AGACTGTACCCATCCAGCCGCTGCACATCACACATTTACTTTAACAGAACAGTCACCGAC 1464  
Qy 841 AACATGCTGTGTGTGAGACACTCGGAGCGCGGCGGCCCAAGCAAACTTGCACGACGCC 900  
Db 1465 AACATGCTGTGTGTGAGACACTCGGAGCGCGGCGGCCCAAGCAAACTTGCACGACGCC 1524  
Qy 901 TGCAGGCGGATTCGGGAGGCCCTCGTGTGTCTGAACGATGTCGCCGCTGTACACAAAG 960  
Db 1525 TGCAGGCGGATTCGGGAGGCCCTCGTGTGTCTGAACGATGTCGCCGCTGTACACAAAG 1584  
Qy 961 GGCATCATCAGCTGGGCTGGGCTGTGAGACAGAGGATGTCCCGGCTGTACACAAAG 1020  
Db 1585 GGCATCATCAGCTGGGCTGGGCTGTGAGACAGAGGATGTCCCGGCTGTACACAAAG 1644  
Qy 1021 GTTACCAACTACTAGATGGAATTCGTGACAACTGCAACCGTGA 1065  
Db 1645 GTTACCAACTACTAGATGGAATTCGTGACAACTGCAACCGTGA 1689

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 18, 2003, 19:51:53 ; Search time 375.192 Seconds  
(without alignments)  
7662.485 Million cell updates/sec

Title: US-09-987-455-4

Perfect score: 1065

Sequence: 1 tctgagggaacagtgactg.....gtgacaacatgcgaccgtga.1065

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N\_Geneseq\_19Jun03.\*  
1: /SIDSI/gcgdata/geneseq/geneseqn-embl/NA1980.DAT.\*  
2: /SIDSI/gcgdata/geneseq/geneseqn-embl/NA1981.DAT.\*  
3: /SIDSI/gcgdata/geneseq/geneseqn-embl/NA1982.DAT.\*  
4: /SIDSI/gcgdata/geneseq/geneseqn-embl/NA1983.DAT.\*  
5: /SIDSI/gcgdata/geneseq/geneseqn-embl/NA1984.DAT.\*  
6: /SIDSI/gcgdata/geneseq/geneseqn-embl/NA1985.DAT.\*  
7: /SIDSI/gcgdata/geneseq/geneseqn-embl/NA1986.DAT.\*  
8: /SIDSI/gcgdata/geneseq/geneseqn-embl/NA1987.DAT.\*  
9: /SIDSI/gcgdata/geneseq/geneseqn-embl/NA1988.DAT.\*  
10: /SIDSI/gcgdata/geneseq/geneseqn-embl/NA1989.DAT.\*  
11: /SIDSI/gcgdata/geneseq/geneseqn-embl/NA1990.DAT.\*  
12: /SIDSI/gcgdata/geneseq/geneseqn-embl/NA1991.DAT.\*  
13: /SIDSI/gcgdata/geneseq/geneseqn-embl/NA1992.DAT.\*  
14: /SIDSI/gcgdata/geneseq/geneseqn-embl/NA1993.DAT.\*  
15: /SIDSI/gcgdata/geneseq/geneseqn-embl/NA1994.DAT.\*  
16: /SIDSI/gcgdata/geneseq/geneseqn-embl/NA1995.DAT.\*  
17: /SIDSI/gcgdata/geneseq/geneseqn-embl/NA1996.DAT.\*  
18: /SIDSI/gcgdata/geneseq/geneseqn-embl/NA1997.DAT.\*  
19: /SIDSI/gcgdata/geneseq/geneseqn-embl/NA1998.DAT.\*  
20: /SIDSI/gcgdata/geneseq/geneseqn-embl/NA1999.DAT.\*  
21: /SIDSI/gcgdata/geneseq/geneseqn-embl/NA2000.DAT.\*  
22: /SIDSI/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT.\*  
23: /SIDSI/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT.\*  
24: /SIDSI/gcgdata/geneseq/geneseqn-embl/NA2002.DAT.\*  
25: /SIDSI/gcgdata/geneseq/geneseqn-embl/NA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1065	100.0	1065	24	Human kringle 2 se
2	1065	100.0	1068	10	Sequence of coding
3	1065	100.0	1170	10	Sequence of coding
4	1065	100.0	1314	10	Sequence of coding
5	1065	100.0	1419	10	Sequence of coding
6	1065	100.0	1680	17	Novel plasminogen
7	1065	100.0	1680	17	Novel plasminogen
8	1065	100.0	1689	9	Tissue plasminogen

9	1065	100.0	1780	11	AAQ01358	Sequence encoding
10	1065	100.0	2100	10	AAQ01119	Sequence of native
11	1065	100.0	2100	11	AAQ05532	Plasmid pST112 enc
12	1065	100.0	2162	16	AAQ08730	Human tPA clone.
13	1065	100.0	2162	19	AAV37294	Human tissue plas
14	1065	100.0	2547	7	AAQ06059	Sequence encoding
15	1065	100.0	2560	10	AAQ05042	Plasmid pKG12 cont
16	1065	100.0	7533	8	AAQ070491	Entire sequence of
17	1065	100.0	7533	9	AAQ081970	Plasmid pEMpl-tPA.
18	1063.4	99.8	1614	11	AAQ03581	Glycosylated plas
19	1063.4	99.8	1614	12	AAQ11550	Sequence encoding
20	1063.4	99.8	1659	12	AAQ11551	Sequence encoding
21	1063.4	99.8	1689	9	AAQ082179	Tissue plasminogen
22	1063.4	99.8	1689	9	AAQ081090	Mutated t-PA analo
23	1063.4	99.8	1689	11	AAQ04903	Part of tPA024 gen
24	1063.4	99.8	1689	11	AAQ04904	Part of tPA023 gen
25	1063.4	99.8	1689	11	AAQ05177	Sequence encoding
26	1063.4	99.8	1689	12	AAQ13379	T-PA67* mutant wit
27	1063.4	99.8	1689	24	ABK52546	Human tissue plas
28	1063.4	99.8	1740	17	AAQ27585	Full-length tissue
29	1063.4	99.8	1955	21	AAQ23198	Human tissue plas
30	1063.4	99.8	1989	7	AAQ07000	Sequence encoding
31	1063.4	99.8	2091	12	AAQ12073	T-PA variant contg
32	1063.4	99.8	2099	11	AAQ05534	Plasmid pLGTQTPA
33	1063.4	99.8	2100	12	AAQ12072	T-PA with -ve char
34	1063.4	99.8	2170	8	AAQ070248	Sequence encoding
35	1063.4	99.8	2170	16	AAQ086576	Human tissue plas
36	1063.4	99.8	2249	11	AAQ05535	Plasmid pLGN Delta
37	1063.4	99.8	2457	7	AAQ06146	cDNA sequence enco
38	1063.4	99.8	2509	10	AAQ05058	cDNA of human tiss
39	1063.4	99.8	2519	24	ABK52545	Human tissue plas
40	1063.4	99.8	2519	24	ABK59544	Human tissue type
41	1063.4	99.8	2547	4	AAQ30001	Sequence encoding
42	1063.4	99.8	7360	17	AAQ15930	DHFR/intron (WTras
43	1062	99.7	1128	24	AAQ0614	OmpA-K2S fusion pr
44	1062	99.7	1341	11	AAQ06762	Sequence encoding
45	1061.8	99.7	1068	10	AAQ01122	Sequence of coding

ALIGNMENTS

RESULT 1  
AAD40616  
ID AAD40616 standard; DNA; 1065 BP.

XX	AC	AAD40616;	
XX	AC	AAD40616;	
XX	DT	30-OCT-2002 (first entry)	
XX	DE	Human kringle 2 serine protease (K2S) DNA.	
XX	XX	Tissue plasminogen activator; tPA; kringle 2 serine protease; stroke;	
KW	KW	K2S; cardiac infarction; acute myocardial infarction; pulmonary embolism;	
KW	KW	artery occlusion; deep vein thrombosis; blood clotting; thrombolytic;	
KW	KW	cerebroprotective; cardiant; ompA; enzyme; human; gene; ds.	
XX	OS	Homo sapiens.	
XX	XX	Key Location/Qualifiers	
FT	FT	CDS 1..1065	
FT	FT	/*tag= a	
FT	FT	/product= "Human K2S protein"	
XX	XX	W0200240650-A2.	
XX	XX	23-MAY-2002.	
XX	XX	07-NOV-2001; 2001WO-BP12857.	
XX	XX	14-NOV-2000; 2000GB-0027779.	
XX	XX	(BOEH ) BOEHRINGER INGELHEIM INT GMBH.	

XX Werner R, Goetz F, Tayapiwatana C, Manosroi J, Manosroi A;  
 XX WPI; 2002-519376/55.  
 DR P-PSDB; AAE25036.  
 XX  
 PT Producing active, correctly folded recombinant tissue plasminogen  
 PT activator, Kingle 2 serine protease in prokaryotic cells by expressing  
 PT the protein-encoding DNA operably linked to DNA coding for signal  
 PT peptide OmpA  
 XX  
 XX Claim 12; Page 31; 80pp; English.  
 XX  
 PS The present invention relates to a method of producing extracellularly  
 CC secreted, active, correctly folded, recombinant tissue plasminogen  
 CC activator (tPA), Kingle 2 serine protease molecule (K2S) or their  
 CC variants in prokaryotic cells by expressing the protein-encoding DNA  
 CC operably linked to DNA coding for signal peptide OmpA. The method is  
 CC useful for producing recombinant DNA-derived tPA, K2S or their variants.  
 CC Sequences of the invention are useful for manufacturing a medicament  
 CC for treating stroke, cardiac infarction, acute myocardial infarction,  
 CC pulmonary embolism, any artery occlusion such as intracranial artery  
 CC occlusion (e.g., arteries supplying the brain), peripherally occluded  
 CC arteries, coronary artery occlusion, deep vein thrombosis or related  
 CC diseases associated with unwanted blood clotting. The present sequence  
 CC is human K2S DNA.  
 XX  
 SQ Sequence 1065 BP; 225 A; 314 C; 312 G; 214 T; 0 other;

Query Match 100.0%; Score 1065; DB 24; Length 1065;  
 Best Local Similarity 100.0%; Pred. No. 9.7e-237;  
 Matches 1065; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCTGAGGAAACAGTACTGCTACTTTGGGAATGGGTGAGCTACCGTGGGACGACAGC 60  
 DB 1 TCTGAGGAAACAGTACTGCTACTTTGGGAATGGGTGAGCTACCGTGGGACGACAGC 60  
 QY 61 CTCACCGAGTGGGTGCTCTCCCTCCCTGGGAATTCATGATCCATAGGCAAGGTT 120  
 DB 61 CTCACCGAGTGGGTGCTCTCCCTCCCTGGGAATTCATGATCCATAGGCAAGGTT 120  
 QY 121 TACACAGACAGAAACCCAGTGGCCAGGACCTGGGCTGGGCAACAATAATTACTGCCGG 180  
 DB 121 TACACAGACAGAAACCCAGTGGCCAGGACCTGGGCTGGGCAACAATAATTACTGCCGG 180  
 QY 181 AATCTGATGGGATGCCAGCCCTGCTGCTCCACCTGGCCCTGAGACCGGAGCTGACGTGG 240  
 DB 181 AATCTGATGGGATGCCAGCCCTGCTGCTCCACCTGGCCCTGAGACCGGAGCTGACGTGG 240  
 QY 241 GAGTACTGTGATGTCCTCTGCTCCACCTGGCCCTGAGACAGTACAGCCAGCTTCAG 300  
 DB 241 GAGTACTGTGATGTCCTCTGCTCCACCTGGCCCTGAGACAGTACAGCCAGCTTCAG 300  
 QY 301 TTTGCGATCAAGGAGGGCTCTTCCGCGACATCGCTCCACCCCTGGCAGGCTGCCATC 360  
 DB 301 TTTGCGATCAAGGAGGGCTCTTCCGCGACATCGCTCCACCCCTGGCAGGCTGCCATC 360  
 QY 361 TTTGCGACGACAGAGGTGCGCCCGGAGAGCGGTTCTGTCGGGGGATATCTCATCAGC 420  
 DB 361 TTTGCGACGACAGAGGTGCGCCCGGAGAGCGGTTCTGTCGGGGGATATCTCATCAGC 420  
 QY 421 TCCTCTGATTTCTCTCTGCGCCCACTGCTTCCAGGAGAGGTTTCCGCCCCACCACTG 480  
 DB 421 TCCTCTGATTTCTCTCTGCGCCCACTGCTTCCAGGAGAGGTTTCCGCCCCACCACTG 480  
 QY 481 ACGGTGATCTTTGGGAGAACATACCGGGTGGTCCCTGGCGAGGAGGACGAAATTTGAA 540  
 DB 481 ACGGTGATCTTTGGGAGAACATACCGGGTGGTCCCTGGCGAGGAGGACGAAATTTGAA 540  
 QY 541 GTCGAAAATACATTTGTCATAGGAATTCGATGATGACCTTACGCAATGACATTCGG 600  
 DB 541 GTCGAAAATACATTTGTCATAGGAATTCGATGATGACCTTACGCAATGACATTCGG 600

QY 601 CTGTGCGAGTGAATCGGATTCGTCCCGTGTGCCCGAGGAGACGCTGGTCCGCACT 660  
 DB 601 CTGTGCGAGTGAATCGGATTCGTCCCGTGTGCCCGAGGAGACGCTGGTCCGCACT 660  
 QY 661 GTGTGCTTCCCGCGGACCTGCGAGCTGCGGACCTGGAACGAGTGTGAGCTCTCCGGC 720  
 DB 661 GTGTGCTTCCCGCGGACCTGCGAGCTGCGGACCTGGAACGAGTGTGAGCTCTCCGGC 720  
 QY 721 TACGGCAAGCATGAGGCGCTTGTCTCTTCTATTCGGAGCGGCTGAAGGAGGCTCATGTC 780  
 DB 721 TACGGCAAGCATGAGGCGCTTGTCTCTTCTATTCGGAGCGGCTGAAGGAGGCTCATGTC 780  
 QY 781 AGACTGTACCATCCAGCCGCTGCACATCAACAATTTACTTTACAGAAAGTCAACGAC 840  
 DB 781 AGACTGTACCATCCAGCCGCTGCACATCAACAATTTACTTTACAGAAAGTCAACGAC 840  
 QY 841 AACATGCTGTGTGTGAGACACTCGGAGCGGCGGCGCCAGGCAAACTTTCACAGCGCC 900  
 DB 841 AACATGCTGTGTGTGAGACACTCGGAGCGGCGGCGGCGCCAGGCAAACTTTCACAGCGCC 900  
 QY 901 TGGCAGGCGCATTCGGGAGCGCCCTGGTGTGTCTGAAAGTGGCCGATCTTGGTG 960  
 DB 901 TGGCAGGCGCATTCGGGAGCGCCCTGGTGTGTCTGAAAGTGGCCGATCTTGGTG 960  
 QY 961 GGCATCATCAAGTGGGCGCTGGGCTGTGGACAGAGGATGTCCCGGCTGTACACAAAG 1020  
 DB 961 GGCATCATCAAGTGGGCGCTGGGCTGTGGACAGAGGATGTCCCGGCTGTACACAAAG 1020  
 QY 1021 GTTACCAACTACCTAGACTGGATTTCGTGACAAACATGCGACCGTGA 1065  
 DB 1021 GTTACCAACTACCTAGACTGGATTTCGTGACAAACATGCGACCGTGA 1065

RESULT 2  
 AAN91120  
 ID AAN91120 standard; DNA; 1068 BP.  
 XX AC AAN91120;  
 XX 25-MAR-2003 (updated)  
 DT 03-OCT-2002 (updated)  
 DT 18-JUN-1990 (first entry)  
 XX  
 DE Sequence of coding region in plasmid pST112.  
 XX  
 KW Tissue plasminogen activator; tPA; thrombolytic agent;  
 KW plasminogen; vascular diseases.  
 XX  
 OS Synthetic.  
 XX  
 XX Key Location/Qualifiers  
 FT CDS 1..1065  
 FT /\*tag= a  
 XX  
 XX EP302456-A.  
 XX  
 XX 08-FEB-1989.  
 XX  
 XX 02-AUG-1988; 88EP-0112569.  
 XX  
 XX 03-AUG-1987; 87GB-0018298.  
 XX 26-OCT-1987; 87GB-0025052.  
 XX 13-NOV-1987; 87GB-0026683.  
 XX  
 XX (FUJI ) FUJISAWA PHARM CO LTD.  
 XX  
 XX Niwa M, Saito Y, Sasaki H, Hayashi M, Notani J, Kobayashi M;  
 XX WPI; 1989-040625/06.  
 XX P-PSDB; AAP94407.  
 XX  
 XX New tissue plasminogen activator -  
 XX comprising finger and growth factor domains lacking tPA for  
 PT





QY 121 TACACAGACAGAACCCAGTCCAGGCACTGGGCTGGGCAACATAAATTACTGCCGG 180  
 DB |||||  
 QY 226 TACACAGACAGAACCCAGTCCAGGCACTGGGCTGGGCAACATAAATTACTGCCGG 285  
 DB |||||  
 QY 181 AATCCTGATGGGATGCCAAGCCCTGGTGCCACGCTGTGAAGAACCCGAGCTGACGTGG 240  
 DB |||||  
 QY 286 AATCCTGATGGGATGCCAAGCCCTGGTGCCACGCTGTGAAGAACCCGAGCTGACGTGG 345  
 DB |||||  
 QY 241 GAGTACTGTGATGTCCTCTCTGCTCCACCTGGGCTGGACAGTACAGCCAGCTTCAG 300  
 DB |||||  
 QY 346 GAGTACTGTGATGTCCTCTCTGCTCCACCTGGGCTGGACAGTACAGCCAGCTTCAG 405  
 DB |||||  
 QY 301 TTTCGCATCAAGAGGAGGCTCTTCGCGACATCGCTCCACCCCTGGCAGGCTGCCATC 360  
 DB |||||  
 QY 406 TTTCGCATCAAGAGGAGGCTCTTCGCGACATCGCTCCACCCCTGGCAGGCTGCCATC 465  
 DB |||||  
 QY 361 TTTCGCAAGCAGAGAGTCCGCGAGAGCGGTTCTGTGCGGGGCATATCATCAGC 420  
 DB |||||  
 QY 466 TTTCGCAAGCAGAGAGTCCGCGAGAGCGGTTCTGTGCGGGGCATATCATCAGC 525  
 DB |||||  
 QY 421 TCCTGCTGGATTCCTCTGCGGCCACCTGCTTCAGGAGAGGTTTCGCGGCCACACCTG 480  
 DB |||||  
 QY 526 TCCTGCTGGATTCCTCTGCGGCCACCTGCTTCAGGAGAGGTTTCGCGGCCACACCTG 585  
 DB |||||  
 QY 481 ACGGTGATCTTGGGCGAGAACATACCGGCTGTGCTCCGCGAGGAGCAGAAATTTGAA 540  
 DB |||||  
 QY 586 ACGGTGATCTTGGGCGAGAACATACCGGCTGTGCTCCGCGAGGAGCAGAAATTTGAA 645  
 DB |||||  
 QY 541 GTGCAAAATACATTTGTCATAAGGAATTCGATGATGACACTTACGAAATGACATTCGC 600  
 DB |||||  
 QY 646 GTGCAAAATACATTTGTCATAAGGAATTCGATGATGACACTTACGAAATGACATTCGC 705  
 DB |||||  
 QY 601 CTGCTGACGCTGAATTCGATTCGCTGCTGCGCCAGGAGAGCGTGGTCCGCACT 660  
 DB |||||  
 QY 706 CTGCTGACGCTGAATTCGATTCGCTGCTGCGCCAGGAGAGCGTGGTCCGCACT 765  
 DB |||||  
 QY 661 GTGTGCTCTCCCGCGGACCTGAGCTGCGGAGTGGAGCTGAGCTCTCCGGC 720  
 DB |||||  
 QY 766 GTGTGCTCTCCCGCGGACCTGAGCTGCGGAGTGGAGCTGAGCTCTCCGGC 825  
 DB |||||  
 QY 721 TACGCAAGCATAGGCTTGTCTCTTTCTATTTCGAGCGGCTGAAGAGGCTCATGTC 780  
 DB |||||  
 QY 826 TACGCAAGCATAGGCTTGTCTCTTTCTATTTCGAGCGGCTGAAGAGGCTCATGTC 885  
 DB |||||  
 QY 781 AGACTGTACCATCCAGCGCTGCACATCAACAATTTACTTACAGAACAGTACCGAC 840  
 DB |||||  
 QY 886 AGACTGTACCATCCAGCGCTGCACATCAACAATTTACTTACAGAACAGTACCGAC 945  
 DB |||||  
 QY 841 AACATGCTGTGTGCGAGACACTCGGAGCGGCGGCGCCAGGCAAACTTGCAGAGCC 900  
 DB |||||  
 QY 946 AACATGCTGTGTGCGAGACACTCGGAGCGGCGGCGCCAGGCAAACTTGCAGAGCC 1005  
 DB |||||  
 QY 901 TGCCAGGCGGATTCGGAGGCGCCCTGTGTGTCTGAACGATGGCCGATGCTTTGGTG 960  
 DB |||||  
 QY 1006 TGCCAGGCGGATTCGGAGGCGCCCTGTGTGTCTGAACGATGGCCGATGCTTTGGTG 1065  
 DB |||||  
 QY 961 GGCATCATAGCTGGGCTGGCTGTGGAAGAGATGTCCGGGTGTGTACAAAG 1020  
 DB |||||  
 QY 1066 GGCATCATAGCTGGGCTGGCTGTGGAAGAGATGTCCGGGTGTGTACAAAG 1125  
 DB |||||  
 QY 1021 GTTACCAACTACCTAGATGATTCGTCACAAACATGCGACCGTGA 1065  
 DB |||||  
 QY 1126 GTTACCAACTACCTAGATGATTCGTCACAAACATGCGACCGTGA 1170  
 DB |||||

## RESULT 4

AA91123

ID AA91123 standard; DNA; 1314 BP.

XX

AC AA91123;

XX

DT 25-MAR-2003 (updated)

DT 03-OCT-2002 (updated)  
 DT 18-JUN-1990 (first entry)  
 XX  
 DB Sequence of coding region in plasmid pTQkPA delta trp.  
 XX  
 KW Tissue plasminogen activator; tPA; thrombolytic agent;  
 KW plasminogen; vascular diseases.  
 XX  
 OS Synthetic.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 1..1065  
 FT /\*tag= a  
 XX  
 PN EP302456-A.  
 XX  
 PD 08-FEB-1989.  
 XX  
 PF 02-AUG-1988; 88EP-0112569.  
 XX  
 PR 03-AUG-1987; 87GB-0018298.  
 PR 26-OCT-1987; 87GB-0025052.  
 PR 13-NOV-1987; 87GB-0026683.  
 XX  
 PA (FUJI ) FUJISAWA PHARM CO LTD.  
 XX  
 PI Niwa M, Saito Y, Sasaki H, Hayashi M, Notani J, Kobayashi M;  
 XX  
 DR WPI; 1989-040625/06.  
 DR P-PSDB; AAP94409.  
 XX  
 PT New tissue plasminogen activator -  
 PT comprising finger and growth factor domains lacking tPA for  
 PT longer half-life and stronger thrombolytic activity.  
 XX  
 PS Disclosure; Page ?; 68pp; English.  
 XX  
 CC (Updated on 03-OCT-2002 to add missing OS field.)  
 CC (Updated on 25-MAR-2003 to correct PA field.)  
 XX  
 SQ Sequence 1314 BP; 286 A; 386 C; 393 G; 249 T; 0 other;

Query Match 100.0%; Score 1065; DB 10; Length 1314;  
 Best Local Similarity 100.0%; Pred. No. 1e-236;  
 Matches 1065; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 TCTGAGGAAACAGTGAAGTCTCTGCTCCGCTGGGAAATGGGTACGCTACCGTGGCAGCAGC 60  
 DB 250 TCTGAGGAAACAGTGAAGTCTCTGCTCCGCTGGGAAATGGGTACGCTACCGTGGCAGCAGC 309  
 QY 61 CTCAACGAGTGGGTGCTCTGCTCCGCTGGGAAATGGGTACGCTACCGTGGCAGCAGC 120  
 DB 310 CTCAACGAGTGGGTGCTCTGCTCCGCTGGGAAATGGGTACGCTACCGTGGCAGCAGC 369  
 QY 121 TACACAGCACAGAACCCAGTCCGCGACACTGGGCTGGGCAACATAAATTAATCTGCCGG 180  
 DB 370 TACACAGCACAGAACCCAGTCCGCGACACTGGGCTGGGCAACATAAATTAATCTGCCGG 429  
 QY 181 AATCCTGATGGGATGCCAAGCCCTGGTGCCACGCTGTGAAGAACCCGAGCTGACGTGG 240  
 DB 430 AATCCTGATGGGATGCCAAGCCCTGGTGCCACGCTGTGAAGAACCCGAGCTGACGTGG 489  
 QY 241 GAGTACTGTGATGTCCTCTGCTCCACCTGGGCTGGAGACAGTACAGCCAGCTCAG 300  
 DB 490 GAGTACTGTGATGTCCTCTGCTCCACCTGGGCTGGAGACAGTACAGCCAGCTCAG 549  
 QY 301 TTTTCGCATCAAGAGGAGGCTCTTTCGCGACATCGCTCCCAACCCCTGGCAGGCTGCCATC 360  
 DB 550 TTTTCGCATCAAGAGGAGGCTCTTTCGCGACATCGCTCCCAACCCCTGGCAGGCTGCCATC 609  
 QY 361 TTTTCGCAAGCAGAGGAGGCTGCCCGGAGAGCGGTTCTGTGTGGGGGCATATCATCAGC 420  
 DB 610 TTTTCGCAAGCAGAGGAGGCTGCCCGGAGAGCGGTTCTGTGTGGGGGCATATCATCAGC 669





Db 1456 AACATGCTGTGTGCTGGAGACATCGGAGCGCGCGCCCGCCAGGCAAACTTGCACGACGCC 1515  
Qy 901 TGCCAGGGCGATTCCGGAGGCCCCCTCGTGTGTCTGAACGATGGCGCGATGACTTTGGTG 960  
Db 1516 TGCCAGGGCGATTCCGGAGGCCCCCTCGTGTGTCTGAACGATGGCGCGATGACTTTGGTG 1575  
Qy 961 GGCATCATCAGCTGGGGCTGGGCTGTGTGACAGAAAGGATGTCGGGGTGTGTACAAAG 1020  
Db 1576 GGCATCATCAGCTGGGGCTGGGCTGTGTGACAGAAAGGATGTCGGGGTGTGTACAAAG 1635  
Qy 1021 GTTACCAACTACTAGACTGGATTGCTGACACATGCGACCGTGA 1065  
Db 1636 GTTACCAACTACTAGACTGGATTGCTGACACATGCGACCGTGA 1680

RESULT 7

AAT27587

ID AAT27587 standard; DNA; 1680 BP.

XX AC AAT27587;

XX DT 25-MAR-2003 (updated)

XX DT 06-AUG-1996 (first entry)

XX DE Novel plasminogen activator DNA.

XX KW Tissue plasminogen activator; tPA; alpha2-plasmin inhibitor;

XX KW fibrinolytic; thrombolytic; fibrin; thrombosis; blood clotting;

XX KW protein engineering; kringle; ss.

XX OS Synthetic.

XX PH Key

FT sig\_peptide 1..105

FT mat\_peptide 106..1677

FT /tag= a

FT /tag= b

XX US5504001-A.

XX PN 02-APR-1996.

XX PD 06-JUN-1994; 94US-0254485.

XX PR 25-NOV-1987; 87US-0125629.

XX PR 28-JAN-1992; 92US-0827587.

XX PR 06-JUN-1994; 94US-0254485.

XX PA (ZYMO ) ZYMOGENETICS INC.

XX PI Foster DC;

XX DR WPI; 1996-187699/19.

XX DR P-PSDB; AAR96222.

XX PT Hybrid plasminogen activator comprises human tPA activator and

XX PT N-terminal crosslinking domain from alpha2-plasmin inhibitor

XX PT useful to treat thrombosis and image blood clots

XX PS Example 3; Fig 10; 35pp; English.

XX CC A DNA construct (AAT27587) codes for a novel plasminogen

XX CC activator (AAR96223) in which the kringle K1 domain of plasminogen

XX CC (see also AAR96221) is mutated to substitute asparagine for aspartic

XX CC acid at position 5, and replaces the native K1 domain of tissue

XX CC plasminogen activator (tPA) (AAR96220). It was obtd. by mutagenesis

XX CC of the sequence in vector Zemo9-8000 (see also AAT27587). The novel

XX CC protein can be expressed in Escherichia coli RRI/Zemo9-8010 cells

XX CC (FERM P-9315). Novel plasminogen activators are produced that show

XX CC increased clot lysing specificity or plasma half-life.

XX CC (Updated on 25-MAR-2003 to correct PF field.)

SQ Sequence 1680 BP; 386 A; 471 C; 480 G; 343 T; 0 other;  
Query Match 100.0%; Score 1065; DB 17; Length 1680;  
Best Local Similarity 100.0%; Pred. No. 1.le-236;  
Matches 1065; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 TCTGAGGAAACAGTGTCTACTTTGGGAATGGGTGAGCTACCGTGGCAGCAGCACAGC 60  
Db 616 TCTGAGGAAACAGTGTCTACTTTGGGAATGGGTGAGCTACCGTGGCAGCAGCACAGC 675  
Qy 61 CTCACCGAGTGGGTGCTCTCTGCTCCCGTGGAAATTCATGATCCTGATAGGCAAGTT 120  
Db 676 CTCACCGAGTGGGTGCTCTCTGCTCCCGTGGAAATTCATGATCCTGATAGGCAAGTT 735  
Qy 121 TACACAGCACAGAAACCCAGTGGCCAGGCACTGGGCGCTGGGCAACATAATTTACTGCCGG 180  
Db 736 TACACAGCACAGAAACCCAGTGGCCAGGCACTGGGCGCTGGGCAACATAATTTACTGCCGG 795  
Qy 181 AATCTGATGGGATGCCAAGCCCTGTGTGCCACGTGTGCAAGTGAAGAACCGCAGCGTGAAGTGG 240  
Db 796 AATCTGATGGGATGCCAAGCCCTGTGTGCCACGTGTGCAAGTGAAGAACCGCAGCGTGAAGTGG 855  
Qy 241 GAGTACTGTGATGTCCTCTCTGCTCCACCTGCGGCTGAGACAGTACAGCAGCCTCAG 300  
Db 856 GAGTACTGTGATGTCCTCTCTGCTCCACCTGCGGCTGAGACAGTACAGCAGCCTCAG 915  
Qy 301 TTTGCGCATCAAGAGGAGGCTCTTGGCGGACATCGCTCCACCCCTGGCAGGCTGCCATC 360  
Db 916 TTTGCGCATCAAGAGGAGGCTCTTGGCGGACATCGCTCCACCCCTGGCAGGCTGCCATC 975  
Qy 361 TTTGCCAAGCACAGAGGTCGCCCGGAGAGCGGTTCTGTGCGGGGGCATCTCATACAGC 420  
Db 976 TTTGCCAAGCACAGAGGTCGCCCGGAGAGCGGTTCTGTGCGGGGGCATCTCATACAGC 1035  
Qy 421 TCCTGCTGGATTCTCTGCGCGCCACTGCTTCAGGAGAGGTTTCGCGCCACCAACCTG 480  
Db 1036 TCCTGCTGGATTCTCTGCGCGCCACTGCTTCAGGAGAGGTTTCGCGCCACCAACCTG 1095  
Qy 481 ACGGTGATCTTGGGAGAAACATACCGGTGTCCTCTGGCAGGAGGAGCAGAAATTTGAA 540  
Db 1096 ACGGTGATCTTGGGAGAGNACATACCGGTGTCCTCTGGCAGGAGGAGCAGAAATTTGAA 1155  
Qy 541 GTCGAAAAATACATTTGCTCCATAGGAATTCGATGATGACACTTACGACAAATGACATTCGG 600  
Db 1156 GTCGAAAAATACATTTGCTCCATAGGAATTCGATGATGACACTTACGACAAATGACATTCGG 1215  
Qy 601 CTGCTGCAGCTGAATCGGATTCGTCGCGCTGTCGCCAGGAGAGCAGCGTGGTCCGCACT 660  
Db 1216 CTGCTGCAGCTGAATCGGATTCGTCGCGCTGTCGCCAGGAGAGCAGCGTGGTCCGCACT 1275  
Qy 661 GTGTGCTTCCCGCGGAGCTTCGAGCTGCGCGACTGGAGCGAGTGTGAGCTCTCCGCGC 720  
Db 1276 GTGTGCTTCCCGCGGAGCTTCGAGCTGCGCGACTGGAGCGAGTGTGAGCTCTCCGCGC 1335  
Qy 721 TACGCGAAGCATGAGGCTTGTCTCTTTTATTTCCGAGCGGCTGAAGAGGCTCATGTC 780  
Db 1336 TACGCGAAGCATGAGGCTTGTCTCTTTTATTTCCGAGCGGCTGAAGAGGCTCATGTC 1395  
Qy 781 AGACTGTACCCATCCAGCGCTGCACATCAACAACTTTACTTAACAGAAAGTACCCGAC 840  
Db 1396 AGACTGTACCCATCCAGCGCTGCACATCAACAACTTTACTTAACAGAAAGTACCCGAC 1455  
Qy 841 AACATGCTGTGCTGGAGACACTCGGAGCGCGCGCCCGCCAGGCAAACTTGCACGACGCC 900  
Db 1456 AACATGCTGTGCTGGAGACACTCGGAGCGCGCGCGCCCGCCAGGCAAACTTGCACGACGCC 1515  
Qy 901 TGCCAGGCGGATTCGGGAGGCCCTCTGTGTGTCTGAACGATGGCGCGCATGACTTTGGTG 960  
Db 1516 TGCCAGGCGGATTCGGGAGGCCCTCTGTGTGTCTGAACGATGGCGCGCATGACTTTGGTG 1575  
Qy 961 GGCATCATCAGCTGGGGCTGGGCTGTGTGACAGAAAGGATGTCGGGGTGTGTACAAAG 1020  
Db 1576 GGCATCATCAGCTGGGGCTGGGCTGTGTGACAGAAAGGATGTCGGGGTGTGTACAAAG 1635

QY 1021 GTTACAACTACTAGACTGGATTCTGTCACAACTGCGACCGTGA 1065  
DB 1636 GTTACCACTACTAGACTGGATTCTGTCACAACTGCGACCGTGA 1680

RESULT 8

ID AAN82178  
XX AAN82178 standard; DNA; 1689 BP.

AC AAN82178;

XX 25-MAR-2003 (updated)

DT 10-MAR-2003 (updated)

DT 03-NOV-1990 (first entry)

XX Tissue plasminogen activator with S-119 subetd for M.

DE Tissue plasminogen activator; tPA; thrombosis; N-glycosylation; ss.

XX Unidentified.

XX Key

XX Location/Qualifiers

XX 1..1689

XX /\*tag= a

XX /product=modified tPA

XX JP63230083-A.

XX 26-SEP-1988.

XX 20-MAR-1987; 87JP-0064339.

XX 20-MAR-1987; 87JP-0064339.

XX (EISA ) EISAI CO LTD.

XX WPI; 1988-311961/44.

XX P-PSDB; AAP82581.

XX Modified tissue plasminogen activator - having glycine-183 and serine-186

XX residues sustd. with serine and threonine.

XX Disclosure; Page ?; 16pp; Japanese.

XX One N-glycosylation site, i.e. NSS (117-119) is substituted with

XX NSM and the N-glycosylation is removed.

XX Plasmid encoding the modified tPA is 99-6300 and its transformant is

XX E.coli RRI-Zem 99-6300 (FERM P-9127).

XX This modified tPA, used to treat thrombosis, is of high quality and

XX has a longer half life period in blood.

XX See also AAN82177-N82179.

XX (Updated on 10-MAR-2003 to add missing OS field.)

XX (Updated on 25-MAR-2003 to correct PF field.)

XX (Updated on 25-MAR-2003 to correct PR field.)

XX (Updated on 25-MAR-2003 to correct PA field.)

XX Sequence 1689 BP; 376 A; 482 C; 504 G; 327 T; 0 other;

XX Query Match 100.0%; Score 1065; DB 9; Length 1689;

XX Best Local Similarity 100.0%; Pred. No. 1.1e-236;

XX Matches 1065; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCTGAGGAACAGTACTGCTACTTTGGGAATGGGTGAGCTACCGTGGCAGCAGC 60

DB 625 TCTGAGGAACAGTACTGCTACTTTGGGAATGGGTGAGCTACCGTGGCAGCAGC 684

QY 61 CTCACCGAGTGGGTGCTCTCTGCTCCGCTGGGAATTCATGATCCTGATAGGCAAGTT 120

DB 685 CTCACCGAGTGGGTGCTCTCTGCTCCGCTGGGAATTCATGATCCTGATAGGCAAGTT 744

QY 121 TACACAGCAGAAACCCAGTCCCGAGGCTGGGCTGGGCAACATAATTACTGCCGG 180

DB 745 TACACAGCAGAAACCCAGTCCCGAGGCTGGGCTGGGCAACATAATTACTGCCGG 804  
QY 181 AATCTGATGGGATGCCAAGCCCTGCTGTCACAGTCTGAGAACCCGAGCTGACCTGG 240  
DB 805 AATCTGATGGGATGCCAAGCCCTGCTGTCACAGTCTGAGAACCCGAGCTGACCTGG 864  
QY 241 GAGTACTGTGATGTCCTCTCTGCTCCACCTGCGGCTGAGACAGTACAGCCAGCTCAG 300  
DB 865 GAGTACTGTGATGTCCTCTCTGCTCCACCTGCGGCTGAGACAGTACAGCCAGCTCAG 924  
QY 301 TTTGCGATCAAAAGAGGGCTCTTTCGCGAATGCTCCACCTCCACCTGCGAGCTGCCATC 360  
DB 925 TTTGCGATCAAAAGAGGGCTCTTTCGCGAATGCTCCACCTCCACCTGCGAGCTGCCATC 984  
QY 361 TTTGCGAAGCAGAGGAGTCCCGGAGAGGGTTCCTGTCGGGGGAGTACTCATCAGC 420  
DB 985 TTTGCGAAGCAGAGGAGTCCCGGAGAGGGTTCCTGTCGGGGGAGTACTCATCAGC 1044  
QY 421 TCCTGCTGGATTCTCTCTGCGGCCACTGCTTCCAGGAGAGGTTTCCGCCCCACACCTG 480  
DB 1045 TCCTGCTGGATTCTCTCTGCGGCCACTGCTTCCAGGAGAGGTTTCCGCCCCACACCTG 1104  
QY 481 ACGTGATCTTGGGCAAGAAATACACCGGGTGGTCCCTGCGAGGAGGAGCAAAATTTGAA 540  
DB 1105 ACGTGATCTTGGGCAAGAAATACACCGGGTGGTCCCTGCGAGGAGGAGCAAAATTTGAA 1164  
QY 541 GTCGAAAAATACATTGTCATTAAGGAATTCAGATGATGACACTTACGCAATGATTCGCG 600  
DB 1165 GTCGAAAAATACATTGTCATTAAGGAATTCAGATGATGACACTTACGCAATGATTCGCG 1224  
QY 601 CTGCTGAGCTGAAATCGGATTTCGTCGCCGCTGTCGCCAGAGAGCAGCGTGGTCCGCACT 660  
DB 1225 CTGCTGAGCTGAAATCGGATTTCGTCGCCGCTGTCGCCAGAGAGCAGCGTGGTCCGCACT 1284  
QY 661 GTGTCCTTCCCGCGGAGCTGAGCTGCGGAGTGGACGAGTGGAGTGTGAGCTCTCCGGC 720  
DB 1285 GTGTCCTTCCCGCGGAGCTGAGCTGCGGAGTGGACGAGTGGAGTGTGAGCTCTCCGGC 1344  
QY 721 TAGCGCAAGCATGAGGCTTGTCTCTTCTTATTCGAGCGGCTGGAAGGAGGCTCATGTC 780  
DB 1345 TAGCGCAAGCATGAGGCTTGTCTCTTCTTATTCGAGCGGCTGGAAGGAGGCTCATGTC 1404  
QY 781 AGACTGTACCCATCCAGCCGCTGCACATCAACATTTACTTTACAGAACAGTCAACCGAC 840  
DB 1405 AGACTGTACCCATCCAGCCGCTGCACATCAACATTTACTTTACAGAACAGTCAACCGAC 1464  
QY 841 AACATGCTGTGCTGGAGACACTCGAGCGGGGGCCCCAGGCAAACTTGTGACAGCGCC 900  
DB 1465 AACATGCTGTGCTGGAGACACTCGAGCGGGGGGGCCCCAGGCAAACTTGTGACAGCGCC 1524  
QY 901 TGCAGGGCGATTTCGGAGGCCCCCTGCTGTCGATGATGCGCGCATGACTTTGGTG 960  
DB 1525 TGCAGGGCGATTTCGGAGGCCCCCTGCTGTCGATGATGCGCGCATGACTTTGGTG 1584  
QY 961 GGCATCATCAGTGGGGCTTGGGCTGTGGACAGAGGATGTCCCGGGTGTGTACAAAG 1020  
DB 1585 GGCATCATCAGTGGGGCTTGGGCTGTGGACAGAGGATGTCCCGGGTGTGTACAAAG 1644  
QY 1021 GTTACCAACTACCTAGACTGGATTTCGTGACAAATGCGACCGTGA 1065  
DB 1645 GTTACCAACTACCTAGACTGGATTTCGTGACAAATGCGACCGTGA 1689

RESULT 9

AAQ01358

ID AAQ01358 standard; DNA; 1780 BP.

XX AC AAQ01358;

XX AC AAQ01358;

XX 25-MAR-2003 (updated)

DT 08-SEP-1990 (first entry)

XX Sequence encoding wild type tissue plasminogen activator. (t-PA).

DE

XX Wild type tissue plasminogen activator (t-PA); infarction treatment;  
KW thrombosis treatment; embolism treatment.  
XX Homo sapiens.  
XX Key Location/Qualifiers  
FT misc\_feature 1..123  
FT /\*tag= a  
FT /note="synthetic 5' adaptor"  
FT 1738..1780  
FT /\*tag= b  
FT /note="synthetic 3' adaptor"  
XX EP351246-A.  
XX 17-JAN-1990.  
XX 14-JUL-1989; 89EP-0307194.  
XX 15-JUL-1988; 88DK-0003952.  
XX (NOVO ) NOVO-NORDISK AS.  
XX Petersen LC, Boel E;  
XX WPI; 1990-016567/03.  
XX P-PSDB; AAR04699.  
XX New tissue plasminogen activator (t-PA) analogue -  
PT with higher fibrin selectivity than native t-PA, useful for  
PT treating infarction, thrombosis and embolism  
XX Disclosure; Fig 5A-D; 24pp; English.  
XX It is modified in the patent by replacing one or more codons specifying  
CC a positively charged amino acid. The resulting analogues are inserted  
CC into a replicable expression vector which is used to transform or  
CC transfect a host cell which is grown to express a t-PA analogue. The  
CC t-PA analogue is useful for the treatment of diseases or disorders  
CC associated with the formation of thrombi in blood vessels e.g.  
CC infarctions, thrombosis and embolism. The analogue, in the 1-chain form,  
CC exhibits the properties of a proenzyme; however on plasmin-catalysed  
CC cleavage of the one-chain form, the activity of the 2-chain form is fully  
CC retained. Compared to native t-PA, it has a higher fibrin selectivity,  
CC this results in a fibrinolytic agent with a higher fibrin selectivity,  
CC as the fibrinogenolytic activity induced by the analogue is reduced  
CC relative to that induced by native 1-chain t-PA.  
CC (Updated on 25-MAR-2003 to correct PA field.)  
XX Sequence 1780 BP; 405 A; 513 C; 523 G; 339 T; 0 other;  
Query Match 100.0%; Score 1065; DB 11; Length 1780;  
Best Local Similarity 100.0%; Pred. No. 1.1e-236; Mismatches 0; Indels 0; Gaps 0;  
Matches 1065; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 TCTGAGGAAACAGTGAAGTCTTCTGGGATGGGTGAGTCTGAGTGGGACGACAGC 60  
Db 641 TCTGAGGAAACAGTGAAGTCTTCTGGGATGGGTGAGTCTGAGTGGGACGACAGC 700  
QY 61 CTCACCGAGTGGGTCCTCTGCTCCGCTGGGAATTCATGATCCTGATAGGCAAGTT 120  
Db 701 CTCACCGAGTGGGTCCTCTGCTCCGCTGGGAATTCATGATCCTGATAGGCAAGTT 760  
QY 121 TACACAGACAGAACCCAGTGGCCAGGACCTGGGCTGGGCAACATTAATCTGCGG 180  
Db 761 TACACAGACAGAACCCAGTGGCCAGGACCTGGGCTGGGCAACATTAATCTGCGG 820  
QY 181 AATCTGTATGGGATGCCAAGCCCTGGTCCACAGTCTGAAGAACCGCAGGCTGACGTGG 240  
Db 821 AATCTGTATGGGATGCCAAGCCCTGGTCCACAGTCTGAAGAACCGCAGGCTGACGTGG 880  
QY 241 GAGTACTGTATGTGCCCTCTCTGCTCCACCTGCGGCTGAGACAGTACAGCCAGCCTCAG 300

Db 891 GAGTACTGTATGTGCCCTCTCTGCTCCACCTGCGGCTGAGACAGTACAGCCAGCCTCAG 940  
QY 301 TTTGCGATCAAGAGAGGCTCTTGGCCGACATCGCTCCACCTGCGAGGCTGCCATC 360  
Db 941 TTTGCGATCAAGAGAGGCTCTTGGCCGACATCGCTCCACCTGCGAGGCTGCCATC 1000  
QY 361 TTTGCGAAGCAGAGAGGTCGCCCGGAGAGCGGTTCTGTGCGGGGATATCTCATCAGC 420  
Db 1001 TTTGCGAAGCAGAGAGGTCGCCCGGAGAGCGGTTCTGTGCGGGGATATCTCATCAGC 1060  
QY 421 TCCTGCTGGATCTCTGCGGCCCACTGCTTCCAGGAGAGGTTTCCGCCCCACCACTG 480  
Db 1061 TCCTGCTGGATCTCTGCGGCCCACTGCTTCCAGGAGAGGTTTCCGCCCCACCACTG 1120  
QY 481 ACGGTGATCTTGGGCGAGAAATACCGGGTGTCTCTGCGGAGGAGAGCAGAAATTTGAA 540  
Db 1121 ACGGTGATCTTGGGCGAGAAATACCGGGTGTCTCTGCGGAGGAGAGCAGAAATTTGAA 1180  
QY 541 GTCGAAAATACATTTGTCATTAAGGAATTCATGATGACACTTACGAAATGACATTCG 600  
Db 1181 GTCGAAAATACATTTGTCATTAAGGAATTCATGATGACACTTACGAAATGACATTCG 1240  
QY 601 CTGCTGAGCTGAAATCGGATTCGTCGGCTGTGCGCCAGGAGAGCAGGCTGCTCGC 660  
Db 1241 CTGCTGAGCTGAAATCGGATTCGTCGGCTGTGCGCCAGGAGAGCAGGCTGCTCGC 1300  
QY 661 GTGTGCTTCCCGCGGAGCTGCGAGCTGCGGAGTGGACGAGTGTGAGCTCTCCGCG 720  
Db 1301 GTGTGCTTCCCGCGGAGCTGCGAGCTGCGGAGTGGACGAGTGTGAGCTCTCCGCG 1360  
QY 721 TACGCGAAGCATGAGGCTTGTCTCTTTCTATTTCGGAGCGGCTGAAGGAGGCTCATGTC 780  
Db 1361 TACGCGAAGCATGAGGCTTGTCTCTTTCTATTTCGGAGCGGCTGAAGGAGGCTCATGTC 1420  
QY 781 AGACTGTACCATCCAGCGCTGACATCAACAATTTACTTTACAGAACAGTCAACGAC 840  
Db 1421 AGACTGTACCATCCAGCGCTGACATCAACAATTTACTTTAAACAGAACAGTCAACGAC 1480  
QY 841 AACATGCTGTCTGTGGAGACACTCGGAGCGCGGCGCCAGGCAAACTTGCACGACGCC 900  
Db 1481 AACATGCTGTCTGTGGAGACACTCGGAGCGCGGCGCCAGGCAAACTTGCACGACGCC 1540  
QY 901 TGCCAGGCGGATTCGGGAGGCGCCCTGCTGTGTCTGAACGATGGCCGATGACTTTGGTG 960  
Db 1541 TGCCAGGCGGATTCGGGAGGCGCCCTGCTGTGTCTGAACGATGGCCGATGACTTTGGTG 1600  
QY 961 GGCATCATCAGCTGGGCGCTGGGCTGTGGACAGAGGATGTCCCGGGTGTGTACAAAG 1020  
Db 1601 GGCATCATCAGCTGGGCGCTGGGCTGTGGACAGAGGATGTCCCGGGTGTGTACAAAG 1660  
QY 1021 GTTACCACTACCTAGACTGGATTCGTGACAACTGCGACCGTGA 1065  
Db 1661 GTTACCACTACCTAGACTGGATTCGTGACAACTGCGACCGTGA 1705  
RESULT 10  
AAN91119  
ID AAN91119 standard; DNA; 2100 BP.  
XX AC AAN91119;  
XX AC  
DT 25-MAR-2003 (updated)  
DT 03-OCT-2002 (updated)  
DT 18-JUN-1990 (first entry)  
XX DE Sequence of native tPA in plasmid pST112.  
XX KW Tissue plasminogen activator; tPA; thrombolytic agent;  
XX KW plasminogen; vascular diseases.  
XX OS Synthetic.  
XX XX

FH Key Location/Qualifiers  
FT CDS 25..1710  
XX /\*tag= a  
PN EP302456-A.  
PD 08-FEB-1989.  
XX 02-AUG-1988; 89EP-0112569.  
XX 03-AUG-1987; 87GB-0018298.  
PR 26-OCT-1987; 87GB-0025052.  
PR 13-NOV-1987; 87GB-0026683.  
XX (FUJI ) FUJISAWA PHARM CO LTD.  
XX  
PI Niwa M, Saito Y, Sasaki H, Hayashi M, Notani J, Kobayashi M;  
XX WPI; 1989-040625/06.  
DR P-PSDB; AAP94406.  
XX  
PT New tissue plasminogen activator -  
PT comprising finger and growth factor domains lacking tPA for  
PT longer half-life and stronger thrombolytic activity.  
XX  
PS Example 29; Fig 21; 68pp; English.  
XX  
CC cDNA sequence of native tPA gene is excised from plasmid pStII12, and  
CC digested with BamHI and SalI to form plasmid pStII18.  
CC (Updated on 03-OCT-2002 to add missing OS field.)  
CC (Updated on 25-MAR-2003 to correct PA field.)  
XX  
SQ Sequence 2100 BP; 498 A; 595 C; 597 G; 410 T; 0 other;  
  
Query Match 100.0%; Score 1065; DB 10; Length 2100;  
Best Local Similarity 100.0%; Pred. No. 1.1e-236;  
Matches 1065; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 TCTGAGGAAACAGTACTGCTTCTGGGAATGGTCCAGCTACCGTGGCAGCAGCAGC 60  
DB 649 TCTGAGGAAACAGTACTGCTTCTGGGAATGGTCCAGCTACCGTGGCAGCAGCAGC 708  
QY 61 CTCACCGAGTGGGTCCTCTGCTCCCTCCGCTGGGAATCCATGATCCTGATAGGCAAGTT 120  
DB 709 CTCACCGAGTGGGTCCTCTGCTCCCTCCGCTGGGAATCCATGATCCTGATAGGCAAGTT 768  
QY 121 TACACAGCAGACAGCCAGTCCAGGACCTGGGCTGGGCAACATAATTAATCTGCCGG 180  
DB 769 TACACAGCAGACAGCCAGTCCAGGACCTGGGCTGGGCAACATAATTAATCTGCCGG 828  
QY 181 AATCTGATGGGATGCCAAGCCCTGGTCCAGCTGCTGAGAACCCGAGGCTGACGTGG 240  
DB 829 AATCTGATGGGATGCCAAGCCCTGGTCCAGCTGCTGAGAACCCGAGGCTGACGTGG 888  
QY 241 GAGTACTGTGATGTCCTCTGCTCCACCTGGGCTGGAGACAGTACAGCCAGCCTCAG 300  
DB 889 GAGTACTGTGATGTCCTCTGCTCCACCTGGGCTGGAGACAGTACAGCCAGCCTCAG 948  
QY 301 TTTCGCATCAAGAGGAGGCTTTCGCCGACATGCTCCACCCCTGGCAGGCTGCCATC 360  
DB 949 TTTCGCATCAAGAGGAGGCTTTCGCCGACATGCTCCACCCCTGGCAGGCTGCCATC 1008  
QY 361 TTTCGCAAGCAGAGAGTTCGCCGAGAGGCTTTCGTCGGGGGATCTATCATCAGC 420  
DB 1009 TTTCGCAAGCAGAGAGTTCGCCGAGAGGCTTTCGTCGGGGGATCTATCATCAGC 1068  
QY 421 TCCTGCTGATTCCTCTGCTCCGCCCACTGCTTCCAGGAGAGGTTTCCGCCCCCAACCTG 480  
DB 1069 TCCTGCTGATTCCTCTGCTCCGCCCACTGCTTCCAGGAGAGGTTTCCGCCCCCAACCTG 1128  
QY 481 ACGGTGATCTGGGAGAGATACCGGCTGCTCCCTGGCAGGAGGAGCAGAAATTTGAA 540  
DB 1129 ACGGTGATCTGGGAGAGATACCGGCTGCTCCCTGGCAGGAGGAGCAGAAATTTGAA 1188

QY 541 GTCGAAAAATATCTGTCATAGGAATTCGATGATGACACTTACGACAAATGACATTGCG 600  
DB 1189 GTCGAAAAATATCTGTCATAGGAATTCGATGATGACACTTACGACAAATGACATTGCG 1248  
QY 601 CTGCTGAGCTGAAATCGGATTCGTCGCTGTGTCGCCAGGAGAGCAGCGTGGTCCGCACT 660  
DB 1249 CTGCTGAGCTGAAATCGGATTCGTCGCTGTGTCGCCAGGAGAGCAGCGTGGTCCGCACT 1308  
QY 661 GTGTGCTTCCCGGGGAGCTGACGCTGCGGAGCTGAGCGAGTGTGAGCTCTCCGCG 720  
DB 1309 GTGTGCTTCCCGGGGAGCTGACGCTGCGGAGCTGAGCGAGTGTGAGCTCTCCGCG 1368  
QY 721 TACGCAAGCATGAGGCTTCTCTCTTCTATTCGGAGCGCTGAGGAGGCTCATGTC 780  
DB 1369 TACGCAAGCATGAGGCTTCTCTCTTCTATTCGGAGCGCTGAGGAGGCTCATGTC 1428  
QY 781 AGACTGTACCATCCAGCGCTGACATCAAACTTTTAAACAGAAAGTACCCGAC 840  
DB 1429 AGACTGTACCATCCAGCGCTGACATCAAACTTTTAAACAGAAAGTACCCGAC 1488  
QY 841 AACATGCTGTGCTGGAGACACTCGAGCGCGGCCGCCAGGCAAACTTCGACGACGCC 900  
DB 1489 AACATGCTGTGCTGGAGACACTCGAGCGCGGCCGCCAGGCAAACTTCGACGACGCC 1548  
QY 901 TGCCAGGCGCATTCGGAGGCGCCCTGCTGTGCTGAACGATGGCCGCGCATGCTTGTGTG 960  
DB 1549 TGCCAGGCGCATTCGGAGGCGCCCTGCTGTGCTGAACGATGGCCGCGCATGCTTGTGTG 1608  
QY 961 GGCAATCATGAGTGGGCTGGGCTGTGGACAGAGGATGTCGCGGTTGTACAAAG 1020  
DB 1609 GGCAATCATGAGTGGGCTGGGCTGTGGACAGAGGATGTCGCGGTTGTACAAAG 1668  
QY 1021 GTTACCAACTACCTAGACTGGATTCTGACAACTGCGACCGCTGA 1065  
DB 1669 GTTACCAACTACCTAGACTGGATTCTGACAACTGCGACCGCTGA 1713  
  
RESULT 11  
AAQ05532  
ID AAQ05532 standard; DNA; 2100 BP.  
XX  
AC AAQ05532;  
XX  
DT 25-MAR-2003 (updated)  
DT 11-DEC-1990 (first entry)  
XX  
XX Plasmid pStII12 encoding novel N-terminal for tissue plasminogen  
DE activator (tPA).  
XX  
KW Fibrin; clotting; thrombolytic; vascular disease; stroke;  
KW myocardial infarction; heart attack; pulmonary embolism; ds;  
XX  
OS Homo sapiens.  
XX  
XX Key Location/Qualifiers  
FH CDS 25..1710  
FT /\*tag= a  
FT mat\_peptide 130..1710  
FT /\*tag= b  
XX  
PN EP379890-A.  
XX  
PD 01-AUG-1990.  
XX  
PF 10-JAN-1990; 90EP-0100457.  
XX  
PR 23-JAN-1989; 89GB-0001422.  
XX  
PA (FUJI ) FUJISAWA PHARM CO LTD.  
XX  
XX Niwa M, Satoh S, Suzuki S, Otsuka K, Kusunoki C;  
XX



DR WPI: 1990-232757/31.  
 XX P-PSDB; AAR06237.  
 PT New tissue plasminogen activator - having N-terminal peptide of  
 PT plasminogen linked to tissue plasminogen activator for increased  
 PT stability in vivo.  
 XX  
 PS Disclosure; Page ?; ?pp; English.  
 XX  
 CC tPA with N-terminal peptide of plasminogen linked is more stable in  
 CC vivo than the native form. It is useful as a thrombolytic agent in  
 CC the treatment of vascular diseases eg myocardial infarction,  
 CC pulmonary embolism etc.  
 CC (Updated on 25-MAR-2003 to correct PA field.)  
 XX  
 SQ Sequence 2100 BP; 498 A; 595 C; 597 G; 410 T; 0 other;  
 Query Match 100.0%; Score 1065; DB 11; Length 2100;  
 Best Local Similarity 100.0%; Pred. No. 1.1e-236;  
 Matches 1065; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCTGAGGAAACAGTACTGCTACTTTGGGAATGGGTGAGCTACCGTGGCAACGACACAGC 60  
 Db TCTGAGGAAACAGTACTGCTACTTTGGGAATGGGTGAGCTACCGTGGCAACGACACAGC 708  
 QY 61 CTCACCGAGTGGGTGCTCTCTGCTCCCGTGGGAATTCATGATCCTGATAGGCAAGTT 120  
 Db CTCACCGAGTGGGTGCTCTCTGCTCCCGTGGGAATTCATGATCCTGATAGGCAAGTT 768  
 QY 121 TACACAGCACAGAACCCAGTGCACAGCAGCTGGCCCTGGCCCAACATAATTAATCTGCCG 180  
 Db TACACAGCACAGAACCCAGTGCACAGCAGCTGGCCCTGGCCCAACATAATTAATCTGCCG 828  
 QY 181 AATCTGATGGGATGCCAAGCCCTGTGTGCCACAGTGTGAAGAACCGCAGCTGACGTGG 240  
 Db AATCTGATGGGATGCCAAGCCCTGTGTGCCACAGTGTGAAGAACCGCAGCTGACGTGG 888  
 QY 241 GAGTACTGTGATGTGCTCTCTGCTCCACCTGCGGCTGAGACAGTACAGCCAGCTCAG 300  
 Db GAGTACTGTGATGTGCTCTCTGCTCCACCTGCGGCTGAGACAGTACAGCCAGCTCAG 948  
 QY 301 TTTGCGATCAAGAGGGGCTTTGCGCGACATCGCTCCACCCCTGCGAGCTGCCATC 360  
 Db TTTGCGATCAAGAGGGGCTTTGCGCGACATCGCTCCACCCCTGCGAGCTGCCATC 1008  
 QY 361 TTTGCAAGCACAGAGGTGCGCCGAGAGCGGTTCCTGTGCGGGGATATCTATCAGC 420  
 Db TTTGCAAGCACAGAGGTGCGCCGAGAGCGGTTCCTGTGCGGGGATATCTATCAGC 1068  
 QY 421 TCCTGCTGGATTTCTTCTGCGCCGCCACTGCTTCCAGGAGAGGTTTCCGCCCCACACCTG 480  
 Db TCCTGCTGGATTTCTTCTGCGCCGCCACTGCTTCCAGGAGAGGTTTCCGCCCCACACCTG 1128  
 QY 481 ACGGTGATCTTGGGAGACATACCGGTGTCTCTGCGAGAGGAGGACAGAAATTTGAA 540  
 Db ACGGTGATCTTGGGAGACATACCGGTGTCTCTGCGAGAGGAGGACAGAAATTTGAA 1188  
 QY 541 GTCGAAAATACATGTCTCAATGAAGAAATTCATGATGACACTTACGCAATGACATTCGC 600  
 Db GTCGAAAATACATGTCTCAATGAAGAAATTCATGATGACACTTACGCAATGACATTCGC 1248  
 QY 601 CTGCTGACGCTGAATTCGGATTCGCTCCGCTGTGCCAGGAGAGCGTGTGTCGCACT 660  
 Db CTGCTGACGCTGAATTCGGATTCGCTCCGCTGTGCCAGGAGAGCGTGTGTCGCACT 1308  
 QY 661 GTGTGCTTCCCGCGGAGCTGAGCTGCGGACTGAGAGGATGTGAGCTCTCCGGC 720  
 Db GTGTGCTTCCCGCGGAGCTGAGCTGCGGACTGAGAGGATGTGAGCTCTCCGGC 1368  
 QY 721 TACGGCAAGCATGAGGCTTGTCTCTTCTTATTCGGAGCGGCTGAAGGAGGCTCATGTC 780  
 Db TACGGCAAGCATGAGGCTTGTCTCTTCTTATTCGGAGCGGCTGAAGGAGGCTCATGTC 1428

QY 781 AGACTGTACCATCCAGCGCTGCACATCAACAATTTACTTAACAGAACAGTCAACGAC 840  
 Db AGACTGTACCATCCAGCGCTGCACATCAACAATTTACTTAACAGAACAGTCAACGAC 1488  
 QY 841 AACATGCTGTGTGCTGGAGACACTCGGAGCGGGGCCCCAGGCAAACTTGCACGACGCC 900  
 Db AACATGCTGTGTGCTGGAGACACTCGGAGCGGGGCCCCAGGCAAACTTGCACGACGCC 1548  
 QY 901 TGCCAGGCGGATTCGGAGGCGCCCTGCTGTGTCTGAACGATGGCCGCGATCTTTGGTG 960  
 Db TGCCAGGCGGATTCGGAGGCGCCCTGCTGTGTCTGAACGATGGCCGCGATCTTTGGTG 1608  
 QY 961 GGCATCATCAGCTGGGCGCTGGGCTGTGCGACAGAGGATGTCCCGGGTGTGTACACAAAG 1020  
 Db GGCATCATCAGCTGGGCGCTGGGCTGTGCGACAGAGGATGTCCCGGGTGTGTACACAAAG 1668  
 QY 1021 GTTACCAACTACCTAGACTGGATTCTGTGACAAACATGCGACCGTGA 1065  
 Db GTTACCAACTACCTAGACTGGATTCTGTGACAAACATGCGACCGTGA 1713

RESULT 12  
 AAQ87370  
 ID AAQ87370 standard; DNA; 2162 BP.  
 XX  
 AC AAQ87370;  
 XX  
 DT 25-MAR-2003 (updated)  
 DT 19-SEP-1995 (first entry)  
 XX  
 DE Human tPA clone.  
 XX  
 KW Intron; recombination; combinatorial gene; trans-splicing;  
 KW gene therapy; polymerase chain reaction; PCR; primer; amplification;  
 KW tissue plasminogen activator; tPA; plasmid TPA-KS+; thrombolytic;  
 KW ds.  
 XX  
 OS Synthetic.  
 XX  
 FH Key Location/Qualifiers  
 FT sig\_peptide 82..334 a  
 FT /tag= a  
 FT /note= "signal sequence and finger-like domain"  
 FT CDS 335..447  
 FT /tag= b  
 FT /product= EGF-like domain  
 FT CDS 448..714 c  
 FT /tag= c  
 FT /product= Kringle-1 domain  
 FT CDS 715..972 d  
 FT /tag= d  
 FT /product= Kringle-2 domain  
 FT CDS 973..2162 e  
 FT /tag= e  
 FT /product= catalytic domain  
 XX  
 FN W09507351-A1.  
 XX  
 PD 16-MAR-1995.  
 XX  
 PF 12-SEP-1994; 94WO-US10146.  
 XX  
 PR 10-SEP-1993; 93US-0119512.  
 XX  
 PA (HARD ) HARVARD COLLEGE.  
 XX  
 XX Jarrell KA;  
 XX WPI; 1995-123425/16.  
 DR  
 XX  
 PT New intron-mediated recombinant techniques - used for the  
 PT generation and selection of novel genes and gene prods. for use  
 PT in therapy

XX  
PS  
XX  
CC  
CC  
CC  
CC  
CC  
CC  
CC  
CC  
CC  
SQ

Example 4; Page 59-60; 87pp; English.

A cDNA clone of human tissue plasminogen activator (tPA) was amplified by PCR using the primers given in AAQ87368-69. The amplified tPA DNA (AAQ87370) was ligated into vector KS+ to obtain plasmid TPA-KS+. The construct was used in combinatorial methods involving RNA splicing-mediated shuffling of tPA domains in plasmid pINVI (AAQ87347) to generate novel tPAs having improved thrombolytic properties.

CC (Updated on 25-MAR-2003 to correct PN field.)

Query Match 100.08; Score 1065; DB 16; Length 2162;  
Best Local Similarity 100.08; Pred. No. 1.1e-216;  
Matches 1065; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCTGAGGAAACAGTGAAGTCTTCTGGAATGGGTGAGCTACCGTGGCAGCAGCAGC 60  
DB TCTGAGGAAACAGTGAAGTCTTCTGGAATGGGTGAGCTACCGTGGCAGCAGCAGC 765  
QY 61 CTCACCGAGTGGGTGCTCTCTCCCTCCGTGGAAATTCATGATCCTGATAGGCAAGTT 120  
DB CTCACCGAGTGGGTGCTCTCTCCCTCCGTGGAAATTCATGATCCTGATAGGCAAGTT 825  
QY 121 TACACAGACAGAACCCAGTGCAGCAGCAGTGGGCTGGGCAACATAATTAATCTGCGG 180  
DB TACACAGACAGAACCCAGTGCAGCAGCAGTGGGCTGGGCAACATAATTAATCTGCGG 885  
QY 181 AATCTGATGGGATGCCAAGCCCTGTGTGCACAGTGTGTAAGAACCGCAGGCTGACGTGG 240  
DB AATCTGATGGGATGCCAAGCCCTGTGTGCACAGTGTGTAAGAACCGCAGGCTGACGTGG 945  
QY 241 GAGTACTGTGATGTCCTCTCTGCTCCACCTGCGGCTGGAGACAGTACAGCCAGCTCAG 300  
DB GAGTACTGTGATGTCCTCTCTGCTCCACCTGCGGCTGGAGACAGTACAGCCAGCTCAG 1005  
QY 301 TTTCGCATCAAGAGGAGGCTTTCGCCACATCGCTCCACCCCTGGCAGGCTGCCATC 360  
DB TTTCGCATCAAGAGGAGGCTTTCGCCACATCGCTCCACCCCTGGCAGGCTGCCATC 1065  
QY 361 TTTCGCATCAAGAGGAGGCTTTCGCCACATCGCTCCACCCCTGGCAGGCTGCCATC 420  
DB TTTCGCATCAAGAGGAGGCTTTCGCCACATCGCTCCACCCCTGGCAGGCTGCCATC 1125  
QY 421 TCCTGCTGATGATCTCTCTGCGGCGCCACCTGCTTCAGAGAGAGGTTTCGCCGCCACCTG 480  
DB TCCTGCTGATGATCTCTCTGCGGCGCCACCTGCTTCAGAGAGAGGTTTCGCCGCCACCTG 1185  
QY 481 ACGGTGATCTGGGAGAGACATACCGGCTGCTCCCTGGCAGGAGGAGCAGAAATTTGAA 540  
DB ACGGTGATCTGGGAGAGACATACCGGCTGCTCCCTGGCAGGAGGAGCAGAAATTTGAA 1245  
QY 541 CTCGAAATATCATTTGTCATTAAGGAATTCGATGATGACACTTACGACAAATGACATTCGC 600  
DB CTCGAAATATCATTTGTCATTAAGGAATTCGATGATGACACTTACGACAAATGACATTCGC 1305  
QY 601 CTGCTGAGCTGAATTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 660  
DB CTGCTGAGCTGAATTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1365  
QY 661 GTGTGCTTCCCGGCGGAGCTGAGCTGCGGAGCTGCGGAGCTGCGGAGCTGCGGAGCT 720  
DB GTGTGCTTCCCGGCGGAGCTGAGCTGCGGAGCTGCGGAGCTGCGGAGCTGCGGAGCT 1425  
QY 721 TACGCAAGCATGAGGCTTGTCTCTTTTATTCGAGCGGCTGAAGAGGCTCATGTC 780  
DB TACGCAAGCATGAGGCTTGTCTCTTTTATTCGAGCGGCTGAAGAGGCTCATGTC 1485  
QY 781 AGACTGTACCATTCAGCGGCTGACATCAACATTTACTTAACAGAACAGTCAACGAC 840  
DB AGACTGTACCATTCAGCGGCTGACATCAACATTTACTTAACAGAACAGTCAACGAC 1545

QY 841 AACATGCTGTGCTGCTGAGACACTCGGAGCGGGGCCCCAGGCAAACTTGCACGACGCC 900  
DB AACATGCTGTGCTGCTGAGACACTCGGAGCGGGGCCCCAGGCAAACTTGCACGACGCC 1605  
QY 901 TGCAGGCGGATTCGGGAGGCCCCCTGTGTGCTGTAACGATGCGCGCATGTTGTGTG 960  
DB TGCAGGCGGATTCGGGAGGCCCCCTGTGTGCTGTAACGATGCGCGCATGTTGTGTG 1665  
QY 961 GGCATCATCAGCTGGGCGCTGGGCTGTGGACAGAGGATGTCGCGGTGTGTACAAAG 1020  
DB GGCATCATCAGCTGGGCGCTGGGCTGTGGACAGAGGATGTCGCGGTGTGTACAAAG 1725  
QY 1021 GTTACCAACTACCTAGACTGATTCGTGACAAATGCGACCGTGA 1065  
DB GTTACCAACTACCTAGACTGATTCGTGACAAATGCGACCGTGA 1770

RESULT 13  
AAV37294  
ID AAV37294 standard; DNA; 2162 BP.  
XX AAV37294;  
AC AAV37294;  
XX  
DT 10-SEP-1998 (first entry)  
XX Human tissue plasminogen activator gene sequence.  
DE  
XX  
XX Plasmid pINVI; reverse-splicing intron; group II intron;  
KW exon binding site; domain V motif; branch site acceptor;  
KW nucleophilic group; transesterification; phosphodiester bond;  
KW autocatalytic Y-branched intron; reverse splicing reaction; ds.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT misc\_feature 82..334  
FT /tag= a  
FT /note= "signal sequence and finger-like domain"  
FT misc\_feature 335..447  
FT /tag= b  
FT /note= "EGF-like domain"  
FT misc\_feature 448..714  
FT /tag= c  
FT /note= "Kxingle-1 domain"  
FT misc\_feature 715..972  
FT /tag= d  
FT /note= "Kxingle-2 domain"  
FT misc\_feature 973..2162  
FT /tag= e  
FT /note= "catalytic domain"  
XX  
XX US5780272-A.  
XX  
XX 14-JUL-1998.  
XX  
XX 07-JUN-1995; 95US-0488015.  
XX  
XX 10-SEP-1993; 93US-0119512.  
XX  
XX (HARD ) HARVARD COLLEGE.  
XX  
XX Jarrell KA;  
XX  
XX WPI; 1998-413060/35.  
XX  
XX Reverse splicing construct containing fragments of autocatalytic  
PT introns - able to cleave and ligate discontinuous nucleic acid for  
PT generating new genes and e.g. ribozymes, libraries of enzymes and  
PT antibodies  
XX  
PS Example 4; Columns 53-56; 56pp; English.  
XX

CC The present sequence represents the human tissue plasminogen activator  
CC gene. It was used to construct plasmid tPA-KS+, which is used in the  
CC course of the invention. The specification describes a purified  
CC reverse-splicing intron which comprises a segment comprising a  
CC 5'-part of a group II intron, including an exon binding site not  
CC naturally present in the intron and a second segment comprising a  
CC 3'-part of a group II intron, including a domain V motif, a branch  
CC site acceptor, and a nucleophilic group for transesterifying a  
CC phosphodiester bond of an RNA. Together the two segments form an  
CC autocatalytic Y-branched intron which catalyses integration of at least  
CC the first segment into substrate RNA by a reverse splicing reaction  
CC The reverse-splicing introns are used, by specific cleavage and ligation  
CC of discontinuous nucleic acid, to generate new genes and gene products,  
CC e.g. ribozymes (for use in gene therapy or as reagents in DNA  
CC manipulation, e.g. replacements for restriction enzymes) or  
CC immunologically active or signal-transducing proteins such as antibody  
CC and enzyme libraries.

SQ Sequence 2162 BP; 520 A; 606 C; 619 G; 417 T; 0 other;

Query Match 100.0%; Score 1065; DB 19; Length 2162;  
Best Local Similarity 100.0%; Pred. No. 1.1e-236;  
Matches 1065; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCTGAGGAAACAGTACTGCTACTTTGGGAATGGGTGAGCCTACCGTGGACGACAGC 60  
DB 706 TCTGAGGAAACAGTACTGCTACTTTGGGAATGGGTGAGCCTACCGTGGACGACAGC 765

QY 61 CTCACCGAGTGGGTGCTCTCCCTGCTCCGCTGGAATTCATGATCCTGATAGGCAAGTT 120  
DB 766 CTCACCGAGTGGGTGCTCTCCCTGCTCCGCTGGAATTCATGATCCTGATAGGCAAGTT 825

QY 121 TACACGACAGAAACCCAGTSCCCAGGCACTGGGCTGGGCAACATAATTAATCTGCCGG 180  
DB 826 TACACGACAGAAACCCAGTSCCCAGGCACTGGGCTGGGCAACATAATTAATCTGCCGG 885

QY 181 AATCTGTATGGGATGCCAGCCCTGGTGGCCAGTGTGTAAGAACCGGAGGCTGACGTGG 240  
DB 886 AATCTGTATGGGATGCCAGCCCTGGTGGCCAGTGTGTAAGAACCGGAGGCTGACGTGG 945

QY 241 GAGTACTGTGATGCTCCCTCTCTGCTCCACCTGCGGCTGAGACAGTACAGCCAGCTCAG 300  
DB 946 GAGTACTGTGATGCTCCCTCTCTGCTCCACCTGCGGCTGAGACAGTACAGCCAGCTCAG 1005

QY 301 TTTCCGATCAAGAGGAGGCTCTTCCGCGACATCGCTCCACCCCTGGCAGGCTGCCATC 360  
DB 1006 TTTCCGATCAAGAGGAGGCTCTTCCGCGACATCGCTCCACCCCTGGCAGGCTGCCATC 1065

QY 361 TTTGCCAAGCAGAGAGTGGCCCGGAGAGGGTTCTGTGGGGGGCATACTCATCAGC 420  
DB 1066 TTTGCCAAGCAGAGAGTGGCCCGGAGAGGGTTCTGTGGGGGGCATACTCATCAGC 1125

QY 421 TCCTGCTGATTTCTCTGCGGCCACCTGCTTCCAGGAGAGGTTTCCGCCCCACACCTG 480  
DB 1126 TCCTGCTGATTTCTCTGCGGCCACCTGCTTCCAGGAGAGGTTTCCGCCCCACACCTG 1185

QY 481 ACGGTGATCTTGGGAGAACATACCGGGTGTCTCTGCGAGGAGGAGCAAAATTTGAA 540  
DB 1186 ACGGTGATCTTGGGAGAACATACCGGGTGTCTCTGCGAGGAGGAGCAAAATTTGAA 1245

QY 541 GTCGAAATAATACATTGTCCATAAGGAATTCGATGATGACATTACGAAATGACATTGCG 600  
DB 1246 GTCGAAATAATACATTGTCCATAAGGAATTCGATGATGACATTACGAAATGACATTGCG 1305

QY 601 CTGCTGACGCTGAATTCGATTCGCTGCTGCGGAGGAGGAGGAGGCTGTCGCACT 660  
DB 1306 CTGCTGACGCTGAATTCGATTCGCTGCTGCGGAGGAGGAGGAGGCTGTCGCACT 1365

QY 661 GTGTGCTTCCCGCGGAGCTGCTGAGTGGGAGTGGAGTGGAGTCTTCGCGG 720  
DB 1366 GTGTGCTTCCCGCGGAGCTGCTGAGTGGGAGTGGAGTGGAGTCTTCGCGG 1425

QY 721 TACGCAAGCATAGAGCCCTGCTCTCTTCTATTTCGGAGCGGCTGAAGAGGCTCATGTC 780

DB 1426 TACGCAAGCATAGAGCCCTGCTCTCTTCTATTTCGGAGCGGCTGAAGAGGCTCATGTC 1485  
QY 781 AGACTGTATCCCATCCAGCCGCTGCACATCAACAATTTTAAACAGAACAGTCAACCGAC 840  
DB 1486 AGACTGTATCCCATCCAGCCGCTGCACATCAACAATTTTAAACAGAACAGTCAACCGAC 1545  
QY 841 AACATGCTGTGTCTGGAGACACATCGGAGCGCGGCGGCCAGGCAAACTTGCACGACGCC 900  
DB 1546 AACATGCTGTGTCTGGAGACACATCGGAGCGCGGCGGCCAGGCAAACTTGCACGACGCC 1605  
QY 901 TGCAGGCGGATTCGGAGGCGCCCTGCTGTCTGAACGATGCGCGCATGACTTTGGTG 960  
DB 1606 TGCAGGCGGATTCGGAGGCGCCCTGCTGTCTGAACGATGCGCGCATGACTTTGGTG 1665  
QY 961 GGCATCATCAGCTGGGCGCTGGCTGTGGACAGAGGATGTCCTGGGTGTGTACACAAAG 1020  
DB 1666 GGCATCATCAGCTGGGCGCTGGCTGTGGACAGAGGATGTCCTGGGTGTGTACACAAAG 1725

QY 1021 GTTACCAACTACCTAGACTGGATTTCGTGACAAACATGCGACCGTGA 1065  
DB 1726 GTTACCAACTACCTAGACTGGATTTCGTGACAAACATGCGACCGTGA 1770

RESULT 14  
AAN60659  
ID AAN60659 standard; DNA; 2547 BP.  
XX AAN60659;  
AC AAN60659;  
XX 25-MAR-2003 (updated)  
DT 22-JUL-1991 (first entry)  
XX Sequence encoding human-pre-tissue plasminogen activator (pre-t-PA).  
XX Plasminogen conversion; vascular disease therapy; ss.  
XX Homo sapiens.  
FH Key Location/Qualifiers  
FT CDS 87..191  
FT mat\_peptide /\*tag= a  
FT /\*tag= b  
XX GB2173804-A.  
XX 22-OCT-1986.  
XX 21-APR-1986; 86GB-0609683.  
XX 22-APR-1985; 85US-0725468.  
XX 01-APR-1986; 86US-0846697.  
XX 22-APR-1985; 85US-0725468.  
XX 01-APR-1986; 86US-0846697.  
XX (GETH ) GENENTECH INC.  
XX Heyneker HL, Vohar GA;  
XX WPI; 1986-280715/43.  
XX P-PSDB; AAP60790.  
XX New mutant forms of human tissue plasminogen activator- having  
XX higher specific activity and resistance to conversion to two chain  
XX form  
XX Example; Fig 2; 34pp; English.  
XX The patentors claim a novel recombinant human t-PA which is  
XX resistant to specific enzymatic cleavage because it is stabilised by  
XX site-directed mutagenesis at a 2-chain cleavage site. Partic. the  
XX natural Arg at position 275 is replaced by Gly or Glu, or Ile at

CC position 276 is replaced (275 and 276 refer to the posn. of the AAs  
 CC in the mature protein; i.e. AAs 310 and 311 of AAP60790). Also new are  
 CC (1) DNA sequences; (2) expression vectors; and (3) microorganisms  
 CC and cell cultures transformed with these vectors.  
 CC (Updated on 25-MAR-2003 to correct PR field.)  
 XX  
 SQ Sequence 2547 BP; 627 A; 688 C; 691 G; 541 T; 0 other;  
 Query Match 100.0%; Score 1065; DB 7; Length 2547;  
 Best Local Similarity 100.0%; Pred. No. 1.1e-236; Indels 0; Gaps 0;  
 Matches 1065; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCTGAGGAAACAGTGAAGTCTGCTACTTTGGGAATGGGTGAGCTACCGTGGCAACGACACAGC 60  
 DB 711 TCTGAGGAAACAGTGAAGTCTGCTACTTTGGGAATGGGTGAGCTACCGTGGCAACGACACAGC 770

QY 61 CTCACGAGTGGGTGCTCTGCTCTGCTCCGCTGGGAATTCATGATCTCTGATAGGCAAGGTT 120  
 DB 771 CTCACGAGTGGGTGCTCTGCTCTGCTCCGCTGGGAATTCATGATCTCTGATAGGCAAGGTT 830

QY 121 TACACGACACAGAACCCAGTCCAGGACCTGGGCTGGGCAACATTAATTAATCTCCGG 180  
 DB 831 TACACGACACAGAACCCAGTCCAGGACCTGGGCTGGGCAACATTAATTAATCTCCGG 890

QY 181 AATCTGATGGGATGCCAAGCCCTGGTGCCACGCTGCTGAAGAACCGCAGGCTGACGTTGG 240  
 DB 891 AATCTGATGGGATGCCAAGCCCTGGTGCCACGCTGCTGAAGAACCGCAGGCTGACGTTGG 950

QY 241 GAGTACTGTGATGTCCTCTGCTCCAGTCCAGTGGGCTGGAGACATGACGAGCTCAG 300  
 DB 951 GAGTACTGTGATGTCCTCTGCTCCAGTCCAGTGGGCTGGAGACATGACGAGCTCAG 1010

QY 301 TTTCCCATCAAGAGGAGGCTCTTCCGCGACATCGCTCCACCCCTGGCAGGCTGCCATC 360  
 DB 1011 TTTCCCATCAAGAGGAGGCTCTTCCGCGACATCGCTCCACCCCTGGCAGGCTGCCATC 1070

QY 361 TTTGCCAAGCACAGGAGTTCGCCGAGAGCGGTTCTCTGCGGGGGCATACTCATCAGC 420  
 DB 1071 TTTGCCAAGCACAGGAGTTCGCCGAGAGCGGTTCTCTGCGGGGGCATACTCATCAGC 1130

QY 421 TCCTGCTGGATTCCTCTGCGGCCCACTGCTTCCAGGAGAGGTTTCCGCCCAACCATCTG 480  
 DB 1131 TCCTGCTGGATTCCTCTGCGGCCCACTGCTTCCAGGAGAGGTTTCCGCCCAACCATCTG 1190

QY 481 ACGGTGATCTGGGACAGACATACCGGTGCTCCCTGCGGAGGAGGACGAAATTTGAA 540  
 DB 1191 ACGGTGATCTGGGACAGACATACCGGTGCTCCCTGCGGAGGAGGACGAAATTTGAA 1250

QY 541 CTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAAATGACATTGCG 600  
 DB 1251 CTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAAATGACATTGCG 1310

QY 601 CTGCTGACGTGAATTCGGATTCGTCCTGCTGTCGCCAGAGAGAGCGTGGTCGCACT 660  
 DB 1311 CTGCTGACGTGAATTCGGATTCGTCCTGCTGTCGCCAGAGAGAGCGTGGTCGCACT 1370

QY 661 GTGTGCTTCCCGGGGAGCTGAGCTGCGGACTGGAAGGAGTGAAGTCTCCGGC 720  
 DB 1371 GTGTGCTTCCCGGGGAGCTGAGCTGCGGACTGGAAGGAGTGAAGTCTCCGGC 1430

QY 721 TACGCAAGCATGAGGCTTGTCTCTTCTTATTCGAGCGGCTGAAGAGGCTCATGTC 780  
 DB 1431 TACGCAAGCATGAGGCTTGTCTCTTCTTATTCGAGCGGCTGAAGAGGCTCATGTC 1490

QY 781 AGACTGTACCATTCAGCGCTGCAATCAACAATTTACTTTAACAAGACAGTCAACGAC 840  
 DB 1491 AGACTGTACCATTCAGCGCTGCAATCAACAATTTACTTTAACAAGACAGTCAACGAC 1550

QY 841 AACATGCTGTGTCTGGAGACACTCGGAGCGGCGGCCCAAGGCAAACTTCCAGAGGCC 900  
 DB 1551 AACATGCTGTGTCTGGAGACACTCGGAGCGGCGGCCCAAGGCAAACTTCCAGAGGCC 1610

QY 901 TGCCAGGCGGATTCGGGAGGCGCCCTGCTGTGCTGTGAACGATGGCGCGCATGACTTTGGTG 960

DB 1611 TGCCAGGCGGATTCGGAGGCCCCCTGGTGTCTGAACGATGCGCGCATGACTTTGGTG 1670  
 QY 961 GGCAATCATCAGCTGGGCGCTGGGCTGTGGACAGAGATGTCCCGGTGTGTACACAAAG 1020  
 DB 1671 GGCAATCATCAGCTGGGCGCTGGGCTGTGGACAGAGATGTCCCGGTGTGTACACAAAG 1730

QY 1021 GTTACCAACTACCTAGACTGGATTCTGTGACAAACATGCGACCGTGA 1065  
 DB 1731 GTTACCAACTACCTAGACTGGATTCTGTGACAAACATGCGACCGTGA 1775

RESULT 15  
 ID AAN90542  
 XX AAN90542 standard; cDNA; 2560 BP.  
 AC AAN90542;  
 XX  
 DT 25-MAR-2003 (updated)  
 DT 05-JUN-1990 (first entry)  
 XX  
 DB Plasmid pKG12 contg. DNA encoding human melanoma t-PA.  
 XX Human tissue plasminogen-activator gene; pKG12; ss.  
 XX Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT sig\_peptide 102..206  
 FT /\*tag= a  
 FT mat\_peptide ..1787  
 FT /\*tag= b  
 FT /product=t-PA  
 XX  
 XX EP297066-A.  
 XX  
 XX 28-DEC-1988.  
 XX  
 XX 14-JUN-1988; 88EP-0850207.  
 XX  
 XX 18-JUN-1987; 87SE-0002562.  
 XX  
 XX (KABI ) KABIGEN AB.  
 XX  
 XX Pohl G, Hansson L, Loewenadler B;  
 XX  
 XX WPI; 1989-001503/01.  
 DR P-PSDB; AAP93716.  
 XX  
 XX Modified tissue plasminogen-activator - having domains deleted and amino  
 XX acid changes to increase biological half-life and reduce inactivation.  
 XX  
 XX Disclosure; Fig 1; 18pp; English.  
 XX  
 XX The sequence was obtained from cDNA prep. from mRNA extracted from  
 XX Bowes melanoma cells. The cDNA was used to construct a gene library which  
 XX was screened with a partial t-PA cDNA clone. Plasmid pKG12 was isolated  
 XX and shown to comprise the whole coding region for human t-PA as well as  
 XX 102bp 5' flanking, 760 bp 3' flanking DNA and a poly A tail.  
 XX See also AAN91608.  
 XX (Updated on 25-MAR-2003 to correct PI field.)  
 XX  
 XX Sequence 2560 BP; 634 A; 690 C; 689 G; 547 T; 0 other;

Query Match 100.0%; Score 1065; DB 10; Length 2560;  
 Best Local Similarity 100.0%; Pred. No. 1.1e-236;  
 Matches 1065; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCTGAGGAAACAGTGAAGTCTGCTACTTTGGGAATGGGTGAGCTACCGTGGCAACGACACAGC 60  
 DB 726 TCTGAGGAAACAGTGAAGTCTGCTACTTTGGGAATGGGTGAGCTACCGTGGCAACGACACAGC 785

QY 61 CTCACGAGTGGGTGCTCTGCTCTGCTCCGCTGGGAATTCATGATCTCTGATAGGCAAGGTT 120

Db 786 CTACCGAGTCGGGTGCTCTCGCTCCGCTGGAAATCCATGATCCTGATAGGCAAGGTT 845  
Qy 121 TACACAGCACAGAAACCCAGATGCCAGCACTGGGCCCTGGGCAACATAATTAATGCTGCGG 180  
Db 846 TACACAGCACAGAACCCAGTGCCAGGCACTGGGCCCTGGGCAACATAATTAATGCTGCGG 905  
Qy 181 AATCTGATGGGATGCCAAGCCCTGGTGCCACGCTGCTGAAGAAACCGCAGGCTGACGTTG 240  
Db 906 AATCTGATGGGATGCCAAGCCCTGGTGCCACGCTGCTGAAGAAACCGCAGGCTGACGTTG 965  
Qy 241 GAGTACTGTGATGCTGCTCTGCTCCACCTGCGGCTGAGACAGTACAGCAGCCTCAG 300  
Db 966 GAGTACTGTGATGCTGCTCTGCTCCACCTGCGGCTGAGACAGTACAGCAGCCTCAG 1025  
Qy 301 TTTCCATCAAGAGGAGGCTCTTCCGCGACATCGCTCCACCCCTGGCAGGCTGCCATC 360  
Db 1026 TTTCCATCAAGAGGAGGCTCTTCCGCGACATCGCTCCACCCCTGGCAGGCTGCCATC 1085  
Qy 361 TTTGCCAAGCAGGAGTCCGCGGAGAGCGGTTCTGTGCGGGGCACTACTCATCAGC 420  
Db 1086 TTTGCCAAGCAGGAGTCCGCGGAGAGCGGTTCTGTGCGGGGCACTACTCATCAGC 1145  
Qy 421 TCCTGCTGATCTCTGCGGCCACTGCTTCCAGGAGAGGTTTCCGCCCCACCACTG 480  
Db 1146 TCCTGCTGATCTCTGCGGCCACTGCTTCCAGGAGAGGTTTCCGCCCCACCACTG 1205  
Qy 481 ACGGTGATCTTGGGCGAGAAATACCGGGTGGTCCCTGGCGAGGAGGAGCAAAATTTGAA 540  
Db 1206 ACGGTGATCTTGGGCGAGAAATACCGGGTGGTCCCTGGCGAGGAGGAGCAAAATTTGAA 1265  
Qy 541 GTCGAAATAATGATGCTCAAGAAATTCGATGATGACACTTACGACAAATGACATTCG 600  
Db 1266 GTCGAAATAATGATGCTCAAGAAATTCGATGATGACACTTACGACAAATGACATTCG 1325  
Qy 601 CTGCTGAGCTGAATCGGATTCGCTGCGCTGTCGCCAGGAGAGCGGTGTCGCACT 660  
Db 1326 CTGCTGAGCTGAATCGGATTCGCTGCGCTGTCGCCAGGAGAGCGGTGTCGCACT 1385  
Qy 661 GTGTGCTTCCCGCGGAGCTGCTGCTGCGGAGTGGAGCGAGTGTGAGCTCTCCGGC 720  
Db 1386 GTGTGCTTCCCGCGGAGCTGCTGCTGCGGAGTGGAGCGAGTGTGAGCTCTCCGGC 1445  
Qy 721 TAGGCAAGCATGAGGCTTGTCTCTTTCTATTTCGAGCGGCTGAAGGAGGCTCATGTC 780  
Db 1446 TAGGCAAGCATGAGGCTTGTCTCTTTCTATTTCGAGCGGCTGAAGGAGGCTCATGTC 1505  
Qy 781 AGACTGTACCATCCAGCCGCTGCACATCAACAATTTACTTAACAGAACAGTCAACGAC 840  
Db 1506 AGACTGTACCATCCAGCCGCTGCACATCAACAATTTACTTAACAGAACAGTCAACGAC 1565  
Qy 841 AACATGCTGTGTGAGACACTCGGAGCGGCGGCCCGAGGCAAACTTGCACGAGCC 900  
Db 1566 AACATGCTGTGTGAGACACTCGGAGCGGCGGCCCGAGGCAAACTTGCACGAGCC 1625  
Qy 901 TGCAGGCGCATTCGGAGGCGCCCTGCTGTGCTGAACGATGCGCGCATGACTTTGGTG 960  
Db 1626 TGCAGGCGCATTCGGAGGCGCCCTGCTGTGCTGAACGATGCGCGCATGACTTTGGTG 1685  
Qy 961 GGCATCATCAGTGGGCTGGGCTGTGGACAGAGGATGTCGGGTGTGTACACAAAG 1020  
Db 1686 GGCATCATCAGTGGGCTGGGCTGTGGACAGAGGATGTCGGGTGTGTACACAAAG 1745  
Qy 1021 GTTACCAACTACCTAGACTGATTCGTCGACAAATGCGACCGTGA 1065  
Db 1746 GTTACCAACTACCTAGACTGATTCGTCGACAAATGCGACCGTGA 1790

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 18, 2003, 22:51:23 ; Search time 83.3499 Seconds  
(without alignments)  
5639.757 Million cell updates/sec

Title: US-09-987-455-4  
Perfect score: 1065  
Sequence: 1 tctgagggaacagtactg.....gtgacaacatcgaccgtga 1065

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents NA.\*  
1: /cgn2\_6/ptodata/2/ina/5A\_COMB.seq:\*  
2: /cgn2\_6/ptodata/2/ina/5B\_COMB.seq:\*  
3: /cgn2\_6/ptodata/2/ina/6A\_COMB.seq:\*  
4: /cgn2\_6/ptodata/2/ina/6B\_COMB.seq:\*  
5: /cgn2\_6/ptodata/2/ina/PCTUS\_COMB.seq:\*  
6: /cgn2\_6/ptodata/2/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1065	100.0	1068	2	US-08-811-949-44
2	1065	100.0	1419	2	US-08-811-949-62
3	1065	100.0	1848	3	US-08-814-412-10
4	1065	100.0	2101	2	US-08-811-949-42
5	1065	100.0	2162	1	US-08-119-512-3
6	1065	100.0	2162	1	US-08-488-015B-3
7	1065	100.0	2162	1	US-08-488-015B-25
8	1063.4	99.8	1170	2	US-08-811-949-64
9	1063.4	99.8	1314	2	US-08-811-949-48
10	1063.4	99.8	1738	6	5200340-1
11	1063.4	99.8	2457	6	5344773-1
12	1063.4	99.8	7360	1	US-08-286-740-1
13	1063.4	99.8	7360	5	PCT-US95-09576-1
14	1061.8	99.7	1068	2	US-08-811-949-46
15	1061.8	99.7	1314	2	US-08-811-949-50
16	1061.8	99.7	1974	2	US-08-811-949-38
17	1061.8	99.7	1068	2	US-08-811-949-52
18	1057	99.2	1314	2	US-08-811-949-54
19	1056.4	99.2	1068	1	US-08-137-116-2
20	1056.4	99.2	1137	4	US-09-553-498-9
21	1056.2	99.2	1137	4	US-09-618-869-9
22	1055.4	99.1	1170	2	US-08-811-949-66
23	1054.8	99.0	1065	1	US-08-427-640-1
24	1053.8	98.9	1314	2	US-08-811-949-56
25	1050	98.6	1068	6	5223256-3
26	1046.8	98.3	1065	1	US-08-427-640-5
27					Sequence 5, Appli

28	1041	97.7	1068	2	US-08-811-949-58
29	1035.4	97.2	1068	1	US-08-427-640-3
30	1033	97.0	1065	2	US-08-811-949-60
31	953.4	89.5	2497	6	5185259-2
32	924	86.8	1068	1	US-08-427-640-7
33	801	75.2	1163	2	US-08-558-269-5
34	801	75.2	1163	3	US-09-410-882-5
35	626.6	58.8	1727	6	5244676-4
36	625	58.7	1605	6	5244676-1
37	472	44.3	472	2	US-08-811-949-40
38	417	39.2	453	6	5200340-3
39	264.4	24.8	1724	6	5200340-5
40	170.8	16.0	1236	1	US-07-957-039A-7
41	170.8	16.0	1475	4	US-09-643-597-122
42	170.8	16.0	1475	4	US-09-480-884A-122
43	170.8	16.0	1475	4	US-09-542-615A-122
44	170.8	16.0	1475	4	US-09-606-421B-122
45	170.8	16.0	2301	6	5188829-2

ALIGNMENTS

RESULT 1

US-08-811-949-44  
; Sequence 44, Application US/08811949  
; Patent No. 5840533  
; GENERAL INFORMATION:  
; APPLICANT: NIMA, MINEO  
; APPLICANT: SAITO, YOSHIMASA  
; APPLICANT: SASAKI, HITOSHI  
; APPLICANT: HAYASHI, MASAKO  
; APPLICANT: NOTANI, JOUJI  
; APPLICANT: KOBAYASHI, MASAKAZU  
; TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR  
; NUMBER OF SEQUENCES: 67  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
; ADDRESSEE: P.C.  
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400  
; CITY: ARLINGTON  
; STATE: VA  
; COUNTRY: USA  
; ZIP: 22202  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICANT: US/08/811,949  
; FILING DATE: 05-MAR-1997  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: OBLON, NORMAN F.  
; REGISTRATION NUMBER: 24,618  
; REFERENCE/DOCKET NUMBER: 18-966-0  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 703-413-3000  
; TELEFAX: 703-413-2220  
; INFORMATION FOR SEQ ID NO: 44:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1068 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: circular  
; MOLECULE TYPE: DNA (genomic)  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 1..1065  
US-08-811-949-44

Query Match 100.0%; Score 1065; DB 2; Length 1068;

Best Local Similarity 100.0%; Pred. No. 1.4e-291;  
Matches 1065; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCTGAGGAAACAGTGTCTACTTTGGGAATGGGTACGCTACCGTGGCAACGACAGC 60  
Db 4 TCTGAGGAAACAGTGTCTACTTTGGGAATGGGTACGCTACCGTGGCAACGACAGC 63

QY 61 CTCACCGAGTGGGTCTCTCTGCTCCGCTCCGCTGGAAATCCATGATCTGTAGCAAGTT 120  
Db 64 CTCACCGAGTGGGTCTCTCTGCTCCGCTGGAAATCCATGATCTGTAGCAAGTT 123

QY 121 TACACAGCAGAAACCCAGTGTCCAGGCACTGGGCTGGGCAACATAATTTACTGCCGG 180  
Db 124 TACACAGCAGAAACCCAGTGTCCAGGCACTGGGCTGGGCAACATAATTTACTGCCGG 183

QY 181 AATCTGTATGGGATGCCAAGCCCTGTGTCCAGCTGTGAAGAACCGGCTGACGTTG 240  
Db 184 AATCTGTATGGGATGCCAAGCCCTGTGTCCAGCTGTGTGAAGAACCGGCTGACGTTG 243

QY 241 GAGTACTGTGTGCTCTCTGCTCCAGCTCCAGCTGGGCTGAGACAGTACAGCCAGCTCAG 300  
Db 244 GAGTACTGTGTGCTCTCTGCTCCAGCTGGGCTGAGACAGTACAGCCAGCTCAG 303

QY 301 TTTGCGATCAAAAGGAGGCTCTTGGCGGACATCGCTCCAGCCCTGGCAGGCTGCCATC 360  
Db 304 TTTGCGATCAAAAGGAGGCTCTTGGCGGACATCGCTCCAGCCCTGGCAGGCTGCCATC 363

QY 361 TTTGCGAAGCAGAGAGTGTCCGCGGAGAGCGGTCTCTGTGGGGGGGATCTCATCAGC 420  
Db 364 TTTGCGAAGCAGAGAGTGTCCGCGGAGAGCGGTCTCTGTGGGGGGGATCTCATCAGC 423

QY 421 TCCTGCTGATCTCTCTGCGCGCCACTGTCTTCCAGGAGAGGTTTCCGCCCCACCACTG 480  
Db 424 TCCTGCTGATCTCTCTGCGCGCCACTGTCTTCCAGGAGAGGTTTCCGCCCCACCACTG 483

QY 481 ACGGTGATCTTGGCGAAGACATACCGGGTGTCTTGGCGGAGGAGGAGAGAAATTTGAA 540  
Db 484 ACGGTGATCTTGGCGAAGACATACCGGGTGTCTTGGCGGAGGAGGAGAGAAATTTGAA 543

QY 541 GTGCGAAATATCATTTGTCTCAAGAAATTCGATGATGACATTCAGCAATGACATTCGG 600  
Db 544 GTGCGAAATATCATTTGTCTCAAGAAATTCGATGATGACATTCAGCAATGACATTCGG 603

QY 601 CTGCTGCACTCAATCGATTCGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660  
Db 604 CTGCTGCACTCAATCGATTCGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 663

QY 661 GTGTGCTTCTCCCGCGGACCTGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720  
Db 664 GTGTGCTTCTCCCGCGGACCTGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 723

QY 721 TACGCGAAGCATGAGGCTTGTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 780  
Db 724 TACGCGAAGCATGAGGCTTGTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 783

QY 781 AGACTGTACCCATCCAGCGCTGACATCAGAAATTTACTTAAAGAAAGTACAGCCAGC 840  
Db 784 AGACTGTACCCATCCAGCGCTGACATCAGAAATTTACTTAAAGAAAGTACAGCCAGC 843

QY 841 AACATGCTGTGTGTGGAGACATCTCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 900  
Db 844 AACATGCTGTGTGTGGAGACATCTCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 903

QY 901 TGCCAGGGGATTCGGGAGGCG 960  
Db 904 TGCCAGGGGATTCGGGAGGCG 963

QY 961 GGCATCATCAGTGGGCTGTGAGAGAGGATGTCCGGGGTGTGTACACAAG 1020  
Db 964 GGCATCATCAGTGGGCTGTGAGAGAGGATGTCCGGGGTGTGTGTACACAAG 1023

QY 1021 GTTACCAACTACCTAGATGGATTCGTGACAAACATCGACCGCTGA 1065

Db 1024 GTTACCAACTACCTAGATGGATTCGTGACAAACATCGACCGCTGA 1068

RESULT 2  
US-08-811-949-62  
; Sequence 62, Application US/08811949  
; Patent No. 5840533  
; GENERAL INFORMATION:  
; APPLICANT: NIWA, MINEO  
; APPLICANT: SAITO, YOSHIMASA  
; APPLICANT: SASAKI, HITOSHI  
; APPLICANT: HAYASHI, MASAKO  
; APPLICANT: NOTANI, JOUJI  
; APPLICANT: KOBAYASHI, MASAKAZU  
; TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR  
; NUMBER OF SEQUENCES: 67  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
; ADDRESSEE: P.C.  
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400  
; CITY: ARLINGTON  
; STATE: VA  
; COUNTRY: USA  
; ZIP: 22202  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/811,949  
; FILING DATE: 05-MAR-1997  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: OBLON, NORMAN F.  
; REGISTRATION NUMBER: 24,618  
; REFERENCE/DOCKET NUMBER: 18-966-0  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 703-413-3000  
; TELEFAX: 703-413-2220  
; INFORMATION FOR SEQ ID NO: 62:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1419 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: circular  
; MOLECULE TYPE: DNA (genomic)  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 1..1416  
US-08-811-949-62

Query Match 100.0%; Score 1065; DB 2; Length 1419;  
Best Local Similarity 100.0%; Pred. No. 1.5e-291;  
Matches 1065; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCTGAGGAAACAGTGTCTACTTTGGGAATGGGTACGCTACCGTGGCAACGACAGC 60  
Db 355 TCTGAGGAAACAGTGTCTACTTTGGGAATGGGTACGCTACCGTGGCAACGACAGC 414

QY 61 CTCACCGAGTGGGTCTCTCTGCTCCGCTGGAAATCCATGATCTGTAGCAAGTT 120  
Db 415 CTCACCGAGTGGGTCTCTCTGCTCCGCTGGAAATCCATGATCTGTAGCAAGTT 474

QY 121 TACACAGCAGAAACCCAGTGTCCAGGCACTGGGCTGGGCAACATAATTTACTGCCGG 180  
Db 475 TACACAGCAGAAACCCAGTGTCCAGGCACTGGGCTGGGCAACATAATTTACTGCCGG 534

QY 181 AATCTGTATGGGATGCCAAGCCCTGTGTCCAGCTGTGAAGAACCGGCTGACGTTG 240  
Db 535 AATCTGTATGGGATGCCAAGCCCTGTGTCCAGCTGTGTGAAGAACCGGCTGACGTTG 594

QY 241 GAGTACTGTGTGCTCTCTGCTCCAGCTGGGCTGAGACAGTACAGCCAGCTCAG 300





QY 661 GTGTGCTTCCCGCGGAGCTGCTGAGCTGCGGAGCTGGACGAGTGTGAGCTCTCCGGC 720  
DB 1356 GTGTGCTTCCCGCGGAGCTGCTGAGCTGCGGAGCTGGACGAGTGTGAGCTCTCCGGC 1415  
QY 721 TACGCAAGCATGAGGCTTGTCTCTTCTATTTCCGAGCGGCTGAAAGAGGCTCATGTC 780  
DB 1416 TACGCAAGCATGAGGCTTGTCTCTTCTATTTCCGAGCGGCTGAAAGAGGCTCATGTC 1475  
QY 781 AGACTGTACCATCCAGCGGCTGCATCAACAATTTACTTACAGACAGTCAACGAC 840  
DB 1476 AGACTGTACCATCCAGCGGCTGCATCAACAATTTACTTACAGACAGTCAACGAC 1535  
QY 841 AACATGCTGTGTGAGACACTCGGAGCGGCGGCCAGGCAAACTTGCACGAGCC 900  
DB 1536 AACATGCTGTGTGAGACACTCGGAGCGGCGGCCAGGCAAACTTGCACGAGCC 1595  
QY 901 TGCCAGGCGGATTCGGGAGGCGCCCTGTGTGTCTGAACAGATGCGCGCATGTTTGGTG 960  
DB 1596 TGCCAGGCGGATTCGGGAGGCGCCCTGTGTGTCTGAACAGATGCGCGCATGTTTGGTG 1655  
QY 961 GGCATCATGCTGGGCGCTGGGCTGTGACAGAGGATGTCGGGCTGTGTACACAAG 1020  
DB 1656 GGCATCATGCTGGGCGCTGGGCTGTGACAGAGGATGTCGGGCTGTGTACACAAG 1715  
QY 1021 GTTACCACTACCTAGACTGATTCGTGACAACTGCGACCGTGA 1065  
DB 1716 GTTACCACTACCTAGACTGATTCGTGACAACTGCGACCGTGA 1760

## RESULT 4

US-08-811-949-42  
; Sequence 42, Application US/08811949  
; Patent No. 5840533  
; GENERAL INFORMATION:  
; APPLICANT: NIWA, MINEO  
; APPLICANT: SAITO, YOSHIMASA  
; APPLICANT: SASAKI, HITOSHI  
; APPLICANT: HAYASHI, MASAKO  
; APPLICANT: NOTANI, JOUJI  
; APPLICANT: KOBAYASHI, MASAKAZU  
; TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR  
; NUMBER OF SEQUENCES: 67  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,  
; ADDRESSEE: P.C.  
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400  
; CITY: ARLINGTON  
; STATE: VA  
; COUNTRY: USA  
; ZIP: 22202  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/811,949  
; FILING DATE: 05-MAR-1997  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: OBLON, NORMAN F.  
; REGISTRATION NUMBER: 24,618  
; REFERENCE/DOCKET NUMBER: 18-966-0  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 703-413-3000  
; TELEFAX: 703-413-2220  
; INFORMATION FOR SEQ ID NO: 42:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2101 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: circular  
; MOLECULE TYPE: DNA (genomic)

FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 25..1710  
US-08-811-949-42  
Query Match 100.0%; Score 1065; DB 2; Length 2101;  
Best Local Similarity 100.0%; Pred. No. 1.8e-291;  
Matches 1065; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 TCTGAGGGAACAGTACTGCTACTTTGGGAATGGGCTACCGTGGCAGCAGC 60  
DB 649 TCTGAGGGAACAGTACTGCTACTTTGGGAATGGGCTACCGTGGCAGCAGC 708  
QY 61 CTCAACGAGTGGGCTGCTCTCTCCCTCCCGTGGAAATTCATGATCTCTGATAGGCAAGGTT 120  
DB 709 CTCAACGAGTGGGCTGCTCTCTCCCTCCCGTGGAAATTCATGATCTCTGATAGGCAAGGTT 768  
QY 121 TACACAGCACAGAACCCCACTGAGTGGGCTGGGCAAACTAATTAATCTGCCG 180  
DB 769 TACACAGCACAGAACCCCACTGAGTGGGCTGGGCAAACTAATTAATCTGCCG 828  
QY 181 AATCTGTATGGGATGCCAAGCCCTGTGTCAGCTGTGAAGAACCGCAGGCTGACCTGG 240  
DB 829 AATCTGTATGGGATGCCAAGCCCTGTGTCAGCTGTGAAGAACCGCAGGCTGACCTGG 888  
QY 241 GAGTACTGTATGTCCTCTCTGCTCCACCTGGCGCTGAGACAGTACAGCAGCAGCTCAG 300  
DB 889 GAGTACTGTATGTCCTCTCTGCTCCACCTGGCGCTGAGACAGTACAGCAGCAGCTCAG 948  
QY 301 TTTTCGATCAAGAGGAGGCTCTTTGCGCGACATCCCTCTCCACCTGGCGAGGCTGCCATC 360  
DB 949 TTTTCGATCAAGAGGAGGCTCTTTGCGCGACATCCCTCTCCACCTGGCGAGGCTGCCATC 1008  
QY 361 TTTTCGAGCACAGGAGGCTGCGCGGAGAGGCTTCTGTGCGGGGCTACTCTATCAGC 420  
DB 1009 TTTTCGAGCACAGGAGGCTGCGCGGAGAGGCTTCTGTGCGGGGCTACTCTATCAGC 1068  
QY 421 TCTGTGTGATTTCTCTGCGCGCTCTGCTCCAGGAGAGGTTTCCGCGCCCACTCAGCTG 480  
DB 1069 TCTGTGTGATTTCTCTGCGCGCTCTGCTCCAGGAGAGGTTTCCGCGCCCACTCAGCTG 1128  
QY 481 ACGGTGATTTGGGCGAACAATACCGGCTGCTCCCTGGCGAGGAGGAGCAAAATTTGAA 540  
DB 1129 ACGGTGATTTGGGCGAACAATACCGGCTGCTCCCTGGCGAGGAGGAGCAAAATTTGAA 1188  
QY 541 GTCGAAAAATACATTTGCTCCATAGGATTCGATGATGACACTTACGACATGACATTCGC 600  
DB 1189 GTCGAAAAATACATTTGCTCCATAGGATTCGATGATGACACTTACGACATGACATTCGC 1248  
QY 601 CTGCTGAGCTGAAATCGGATTCGTCCTGCTGTCCTCCAGGAGCAGCAGCTGGTCCGCACT 660  
DB 1249 CTGCTGAGCTGAAATCGGATTCGTCCTGCTGTCCTCCAGGAGCAGCAGCTGGTCCGCACT 1308  
QY 661 GTGTGCTTCCCGCGGAGCTGCTGAGTCCCGGAGCTGGAAGGAGTGTGAGCTCTCCGGC 720  
DB 1309 GTGTGCTTCCCGCGGAGCTGCTGAGTCCCGGAGCTGGAAGGAGTGTGAGCTCTCCGGC 1368  
QY 721 TACGGCAAGCATGAGGCTTGTCTCTTTCTATTCGAGCGGCTGAGAGGCTCATGTC 780  
DB 1369 TACGGCAAGCATGAGGCTTGTCTCTTTCTATTCGAGCGGCTGAGAGGCTCATGTC 1428  
QY 781 AGACTGTACCCATCCAGCCCTGCACATCAACAATTTACTTACAGAAAGTCAACCGAC 840  
DB 1429 AGACTGTACCCATCCAGCCCTGCACATCAACAATTTACTTACAGAAAGTCAACCGAC 1488  
QY 841 AACATGCTGTGTGAGACACTCGGAGCGGCGGCCAGGCAAACTTGCACGAGCC 900  
DB 1489 AACATGCTGTGTGAGACACTCGGAGCGGCGGCCAGGCAAACTTGCACGAGCC 1548  
QY 901 TGCAGGCGGATTCGGGAGGCGCCCTGTGTGTCGAACGATGCGCCGCTGATTTGGTG 960  
DB 1549 TGCAGGCGGATTCGGGAGGCGCCCTGTGTGTCGAACGATGCGCCGCTGATTTGGTG 1608

QY 961 GCATCATCAGTGGGGCTGGGCTGTGGACGAGAGATGTCCCGGGTGTGTACACAAAG 1020  
Db 1609 GCATCATCAGTGGGGCTGGGCTGTGGACGAGAGATGTCCCGGGTGTGTACACAAAG 1668  
QY 1021 GTTACCAACTACTAGACTGGATTCTGTCACACATGCGACCGTGA 1065  
Db 1669 GTTACCAACTACTAGACTGGATTCTGTCACACATGCGACCGTGA 1713

RESULT 5

US-08-119-512-3  
; Sequence 3, Application US/08119512  
; Patent No. 5498531  
; GENERAL INFORMATION:  
; APPLICANT: Jarrell, Kevin A.  
; TITLE OF INVENTION: INTRON MEDIATED RECOMBINANT TECHNIQUES  
; TITLE OF INVENTION: AND REAGENTS  
; NUMBER OF SEQUENCES: 3  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: LAHIVE & COCKFIELD  
; STREET: 60 State Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: ASCII (text)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/119,512  
; FILING DATE: 10-SEP-1993  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Vincent, Matthew P.  
; REGISTRATION NUMBER: 36,709  
; REFERENCE/DOCKET NUMBER: HUI-008  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 227-7400  
; TELEFAX: (617) 227-5941  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2162 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: both  
; MOLECULE TYPE: other nucleic acid  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 82..334  
; OTHER INFORMATION: /product= "Signal Sequence and  
; OTHER INFORMATION: Finger-like domain"  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 335..447  
; OTHER INFORMATION: /product= "EGF-like domain"  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 448..714  
; OTHER INFORMATION: /product= "Kringle-1 domain"  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 715..972  
; OTHER INFORMATION: /product= "Kringle-2 domain"  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 973..2162  
; OTHER INFORMATION: /product= "Catalytic domain"  
US-08-119-512-3

Query Match 100.0%; Score 1065; DB 1; Length 2162;  
Best Local Similarity 100.0%; Pred. No. 1.8e-291;

Matches 1065; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 TCTCAGGAAACAGTACTGTCTACTTTTGGGAATGGGTGAGCTTACCGTGGCAGCGACAGC 60  
Db TCTCAGGAAACAGTACTGTCTACTTTTGGGAATGGGTGAGCTTACCGTGGCAGCGACAGC 765  
QY 61 CTACCCGAGTGGGTGCTCTCTCCGTTGGAAATTCATGATCTCTGATAGGCAAGTTT 120  
Db CTACCCGAGTGGGTGCTCTCTCCGTTGGAAATTCATGATCTCTGATAGGCAAGTTT 825  
QY 121 TACACAGCAGACAGACCCAGTGGCCAGGCACTGGGCTGGGCAACACATAATTTACTGCGG 180  
Db TACACAGCAGACAGACCCAGTGGCCAGGCACTGGGCTGGGCAACACATAATTTACTGCGG 885  
QY 181 AATCTGATGGGATGCCAAGCCCTGGTGGCCAGCTGGGCTGGGCAACACATAATTTACTGCGG 240  
Db AATCTGATGGGATGCCAAGCCCTGGTGGCCAGCTGGGCTGGGCAACACATAATTTACTGCGG 945  
QY 241 GAGTACTGTGATGTCCTCTCTCTCCAGTGGCCCTGAGACAGTACAGCAGCCTCAG 300  
Db GAGTACTGTGATGTCCTCTCTCTCCAGTGGCCCTGAGACAGTACAGCAGCCTCAG 1005  
QY 301 TTTCCGATCAAGAGAGGGCTCTTCCGCGACATCGCCTCCACCCCTGGCAGGCTGCCATC 360  
Db TTTCCGATCAAGAGAGGGCTCTTCCGCGACATCGCCTCCACCCCTGGCAGGCTGCCATC 1065  
QY 361 TTTGCCAAGCAGCAGAGGTGGCCGAGAGCGGTTCCTGCGGGGGGCATCTCATCAGC 420  
Db TTTGCCAAGCAGCAGAGGTGGCCGAGAGCGGTTCCTGCGGGGGGCATCTCATCAGC 1125  
QY 421 TCTGCTGGATTCTCTCTGCGGCCACTGTCTTCAGAGAGGTTTCCGCCCCACACACTG 480  
Db TCTGCTGGATTCTCTCTGCGGCCACTGTCTTCAGAGAGGTTTCCGCCCCACACACTG 1185  
QY 481 ACGTGATCTTGGCAGAACATACCGGGTGGTCCCTCGCGAGGAGGAGCAAAATTTGAA 540  
Db ACGTGATCTTGGCAGAACATACCGGGTGGTCCCTCGCGAGGAGGAGCAAAATTTGAA 1245  
QY 541 GTCGAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAAATGACATTGCG 600  
Db GTCGAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAAATGACATTGCG 1305  
QY 601 CTGCTGAGCTGAATCGGATTGTCCTCGTGTGCCAGGAGACAGCGTGGTCCGCACT 660  
Db CTGCTGAGCTGAATCGGATTGTCCTCGTGTGCCAGGAGACAGCGTGGTCCGCACT 1365  
QY 661 GTGTCCTTCCCGCGGACCTGACAGTCCGAGCTGGAGAGTGTGAGCTTCCGCG 720  
Db GTGTCCTTCCCGCGGACCTGACAGTCCGAGCTGGAGAGTGTGAGCTTCCGCG 1425  
QY 721 TACGCGAAGCATGAGGCTTTGTCTCTCTTCTATTTCGGAGCGGCTGAAGGAGGCTCATGTC 780  
Db TACGCGAAGCATGAGGCTTTGTCTCTCTTCTATTTCGGAGCGGCTGAAGGAGGCTCATGTC 1485  
QY 781 AGACTGTACCCATCCAGCCCTGACATCAACATTTACTTAAACAGACAGTCCCGAC 840  
Db AGACTGTACCCATCCAGCCCTGACATCAACATTTACTTAAACAGACAGTCCCGAC 1545  
QY 841 AACATGCTGTGCTGGAGACACTCGAGCGGGGGCCCCAGGCAAACTTTCAGCAGCGCC 900  
Db AACATGCTGTGCTGGAGACACTCGAGCGGGGGCCCCAGGCAAACTTTCAGCAGCGCC 1605  
QY 901 TGCAGGCGGATTGCGGAGGCCCCCTGCTGTGTCTGAACGATGCGCCGATGACTTTGGTG 960  
Db TGCAGGCGGATTGCGGAGGCCCCCTGCTGTGTGTCTGAACGATGCGCCGATGACTTTGGTG 1665  
QY 961 GGCATCATCAGCTGGGCGCTGGGCTGTGGACAGAGGATGTCCGGGTGTGTACACAAAG 1020  
Db GGCATCATCAGCTGGGCGCTGGGCTGTGGACAGAGGATGTCCGGGTGTGTACACAAAG 1725  
QY 1021 GTTACCAACTACTAGACTGGATTCTGTCACACATGCGACCGTGA 1065  
Db GTTACCAACTACTAGACTGGATTCTGTCACACATGCGACCGTGA 1770



APPLICANT: Jarrell, Kevin A.  
TITLE OF INVENTION: INTRON-MEDIATED RECOMBINANT TECHNIQUES  
NUMBER OF SEQUENCES: 29  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Foley, Hoag & Eliot  
STREET: One Post Office Square  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII (text)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/488,015B  
FILING DATE: 07-JUN-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Vincent, Matthew P.  
REGISTRATION NUMBER: 36,709  
REFERENCE/DOCKET NUMBER: HUV-008.02  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 832-1000  
TELEFAX: (617) 832-7000  
INFORMATION FOR SEQ ID NO: 25:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2162 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
US-08-488-015B-25

Query Match 100.0%; Score 1065; DB 1; Length 2162;  
Best Local Similarity 100.0%; Pred. No. 1.8e-291;  
Matches 1065; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCTGAGGAAACAGTACTGCTTCTGGAATGGTCCAGCTACCGTGGCAGCAGCAGC 60  
DB 706 TCTGAGGAAACAGTACTGCTTCTGGAATGGTCCAGCTACCGTGGCAGCAGCAGC 765  
QY 61 CTCACCGAGTGGGTCCTCTGCTCCCTGCGGAAATTCATGATCCTGATAGGCAAGTT 120  
DB 766 CTCACCGAGTGGGTCCTCTGCTCCCTGCGGAAATTCATGATCCTGATAGGCAAGTT 825  
QY 121 TACACAGACAGAACCCAGTCCAGGACCTGGGCTGGGCAACATAATTAATCTGCCGG 180  
DB 826 TACACAGACAGAACCCAGTCCAGGACCTGGGCTGGGCAACATAATTAATCTGCCGG 885  
QY 181 AATCTGTATGGGATGCCAAGCCCTGCTGCGACGCTGGAAGAACCGCAGGCTGACGTGG 240  
DB 886 AATCTGTATGGGATGCCAAGCCCTGCTGCGACGCTGGAAGAACCGCAGGCTGACGTGG 945  
QY 241 GAGTACTGTGATGTCCTCTGCTCCACCTGCGGCTGAGACAGTACAGCAGCCTCAG 300  
DB 946 GAGTACTGTGATGTCCTCTGCTCCACCTGCGGCTGAGACAGTACAGCAGCCTCAG 1005  
QY 301 TTTCGCATCAAGAGAGGCTCTTCGCCGACATGCTCCACCCCTGCGAGGCTGCCATC 360  
DB 1006 TTTCGCATCAAGAGAGGCTCTTCGCCGACATGCTCCACCCCTGCGAGGCTGCCATC 1065  
QY 361 TTTCGCACAGCAGAGGTCGCCCGGAGAGGCTTCTGCGGGGCGATCTCATCAGC 420  
DB 1066 TTTCGCACAGCAGAGGTCGCCCGGAGAGGCTTCTGCGGGGCGATCTCATCAGC 1125  
QY 421 TCCTGCTGGATCTCTCTGCCGCCACCTGCTTCCAGAGAGAGGTTTCCGCCCCACCACTG 480  
DB 1126 TCCTGCTGGATCTCTCTGCCGCCACCTGCTTCCAGAGAGAGGTTTCCGCCCCACCACTG 1185  
QY 481 ACGTGATCTTGGGCGAGAACATACCGGCTGCTCCCTGCGAGGAGGAGAGCAAAATTTGAA 540  
|||||

Db 1186 ACGTGATCTTGGGCGAGAACATACCGGCTGCTCCCTGCGAGGAGGAGAGCAAAATTTGAA 1245  
QY 541 GTCGAAAAATACATTTGTCATAAGGAATTCGATGATGACACTTACGACAAATGACATTCGG 600  
DB 1246 GTCGAAAAATACATTTGTCATAAGGAATTCGATGATGACACTTACGACAAATGACATTCGG 1305  
QY 601 CTGCTGAGCTGAAATCGGATTCCTCCGCTGTGCCCGAGAGAGAGCGTGGTCGCACT 660  
DB 1306 CTGCTGAGCTGAAATCGGATTCCTCCGCTGTGCCCGAGAGAGAGCGTGGTCGCACT 1365  
QY 661 GTGTGCCCTTCCCGCGGAGACCTGACGTGCCGACTGAGCGAGTGTGAGCTCTCCGGC 720  
DB 1366 GTGTGCCCTTCCCGCGGAGACCTGACGTGCCGACTGAGCGAGTGTGAGCTCTCCGGC 1425  
QY 721 TACGGCAAGCATGAGGCTTGTCTCTCTTCTATTTCGGAGCGGCTGAAAGGAGGCTCATGTC 780  
DB 1426 TACGGCAAGCATGAGGCTTGTCTCTCTTCTATTTCGGAGCGGCTGAAAGGAGGCTCATGTC 1485  
QY 781 AGACTGTACCATCCAGCGCTGCATCATCAACATTTTAAACAGAAAGTCAACCGAC 840  
DB 1486 AGACTGTACCATCCAGCGCTGCATCATCAACATTTTAAACAGAAAGTCAACCGAC 1545  
QY 841 AACATGCTGTGCTGGAGACACTCGGAGCGGCGGCCCGAGGCAAACTTTCACAGCGCC 900  
DB 1546 AACATGCTGTGCTGGAGACACTCGGAGCGGCGGCCCGAGGCAAACTTTCACAGCGCC 1605  
QY 901 TGCCAGGCGGATTCGGAGGCGGCCCTGCTGTGTCTGAACGATGCGCGCATGCTTTGTG 960  
DB 1606 TGCCAGGCGGATTCGGAGGCGGCCCTGCTGTGTCTGAACGATGCGCGCATGCTTTGTG 1665  
QY 961 GGCATCATCAGTGGGCGCTGGGCTGTGGACAGAGGATGTCGCGGTGTGTACACAAAG 1020  
DB 1666 GGCATCATCAGTGGGCGCTGGGCTGTGGACAGAGGATGTCGCGGTGTGTACACAAAG 1725  
QY 1021 GTTACCAACTACTAGTGGATTCGTGACAACTGCGCGCTCA 1065  
DB 1726 GTTACCAACTACTAGTGGATTCGTGACAACTGCGCGCTCA 1770

RESULT 8  
US-08-811-949-64  
; Sequence 64, Application US/08811949  
; Patent No. 5840533  
; GENERAL INFORMATION:  
; APPLICANT: NIWA, MINEO  
; APPLICANT: SAITO, YOSHIMASA  
; APPLICANT: SASAKI, HIYOSHI  
; APPLICANT: HAYASHI, MASAKO  
; APPLICANT: NOTANI, JOUJI  
; APPLICANT: KOBAYASHI, MASAKAZU  
; TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR  
; NUMBER OF SEQUENCES: 67  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,  
; ADDRESS: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400  
; CITY: ARLINGTON  
; STATE: VA  
; COUNTRY: USA  
; ZIP: 22202  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/811.949  
; FILING DATE: 05-MAR-1997  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: OBLON, NORMAN F.  
; REGISTRATION NUMBER: 24,618  
; REFERENCE/DOCKET NUMBER: 18-966-0

TELECOMMUNICATION INFORMATION:

TELEPHONE: 703-413-3000  
TELEFAX: 703-413-2220  
INFORMATION FOR SEQ ID NO: 64:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1170 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: circular  
MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..1167  
US-08-811-949-64

Query Match 99.8%; Score 1063.4; DB 2; Length 1170;  
Best Local Similarity 99.9%; Pred. No. 4e-291;  
Matches 1064; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TCTGAGGAAACAGTGACTGCTTCTGGAATGGTCCAGCTACCGTGGCAGCAGCAGC 60  
DB 106 TCTGAGGAAACAGTGACTGCTTCTGGAATGGTCCAGCTACCGTGGCAGCAGCAGC 165  
QY 61 CTCACCGAGTGGGTCCTCTGCTCCCTCCGCTGGAATTCATGATCCTCATAGGCAAGTT 120  
DB 166 CTCACCGAGTGGGTCCTCTGCTCCCTCCGCTGGAATTCATGATCCTCATAGGCAAGTT 225  
QY 121 TACACGACACAGAACCCAGTCCCGCAGCAGTGGGCTGGGCAACATAATTAAGTCCGG 180  
DB 226 TACACGACACAGAACCCAGTCCCGCAGCAGTGGGCTGGGCAACATAATTAAGTCCGG 285  
QY 181 AATCTGATGGGATGCCAAGCCCTGCTGCCAGCTGCTGAAGAACCGCAGCTGACGTGG 240  
DB 286 AATCTGATGGGATGCCAAGCCCTGCTGCCAGCTGCTGAAGAACCGCAGCTGACGTGG 345  
QY 241 GAGTACTGTGATGTCCTCTGCTCCACCTGCTCCAGCTGCTGAGACAGTACAGCAGCTCAG 300  
DB 346 GAGTACTGTGATGTCCTCTGCTCCACCTGCTCCAGCTGCTGAGACAGTACAGCAGCTCAG 405  
QY 301 TTTCCGATCAAGAGGAGGCTCTTCCGACATCGCTCCACCCCTGCGAGGCTGCCATC 360  
DB 406 TTTCCGATCAAGAGGAGGCTCTTCCGACATCGCTCCACCCCTGCGAGGCTGCCATC 465  
QY 361 TTTGCCAAGCAGAGAGTCCGCCGAGAGCGGTTCTGTGCGGGGATATCTATCAGC 420  
DB 466 TTTGCCAAGCAGAGAGTCCGCCGAGAGCGGTTCTGTGCGGGGATATCTATCAGC 525  
QY 421 TCTGCTGATGTTCTCTGCGGCCCACTGCTTCCAGGAGAGGTTTCCGCCCCACCACTG 480  
DB 526 TCTGCTGATGTTCTCTGCGGCCCACTGCTTCCAGGAGAGGTTTCCGCCCCACCACTG 585  
QY 481 ACGTGATCTTGGGAGAACATACCGGTTGCTCCCTGCGAGGAGGAGCAGAAATTTGAA 540  
DB 586 ACGTGATCTTGGGAGAACATACCGGTTGCTCCCTGCGAGGAGGAGCAGAAATTTGAA 645  
QY 541 GTCGAAATATCATTTGTCATAAGAAATTCGATGATGACATTCAGCAATGACATTCGCG 600  
DB 646 GTCGAAATATCATTTGTCATAAGAAATTCGATGATGACATTCAGCAATGACATTCGCG 705  
QY 601 CTGCTGACGTGAATTCGATTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660  
DB 706 CTGCTGACGTGAATTCGATTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 765  
QY 661 GTGTGCTCTCCCGGCGGAGCTGAGCTGCGGAGTGGAGGAGTGGAGTGGAGTGGAGTGG 720  
DB 766 GTGTGCTCTCCCGGCGGAGCTGAGCTGCGGAGTGGAGGAGTGGAGTGGAGTGGAGTGG 825  
QY 721 TACGCAAGCATGAGGCTTGTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 780  
DB 826 TACGCAAGCATGAGGCTTGTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 885  
QY 781 AGACTGTACCCATCCAGCGGCTGCAATCAACATTTACTTAACAGAAACAGTCAACCGAC 840

DB 886 AGACTGTACCCATCCAGCGCTGCACATCAACAATTTACTTAACAGAAACAGTCAACCGAC 945  
QY 841 AACATGCTGTGCTGGAGACACTCGGAGCGGGGGCCCGAGGCAAACTTCGACGAGCC 900  
DB 946 AACATGCTGTGCTGGAGACACTCGGAGCGGGGGCCCGAGGCAAACTTCGACGAGCC 1005  
QY 901 TGCCAGGCGGATTCGGGAGGCCCCCTGCTGTGTCTGAACGATGCGCGCATGCTTTGGTG 960  
DB 1006 TGCCAGGCGGATTCGGGAGGCCCCCTGCTGTGTCTGAACGATGCGCGCATGCTTTGGTG 1065  
QY 961 GGCATCATGAGCTGGGCGCTGGGCTGTGAGAGGATGTCGGGGTGTGTACACAAAG 1020  
DB 1066 GGCATCATGAGCTGGGCGCTGGGCTGTGAGAGGATGTCGGGGTGTGTACACAAAG 1125  
QY 1021 GTTACCAACTACCTAGACTGATTCGTGACAAACATGCGCGTGA 1065  
DB 1126 GTTACCAACTACCTAGACTGATTCGTGACAAACATGCGCGTGA 1170

RESULT 9

US-08-811-949-48  
; Sequence 48, Application US/08811949  
; Patent No. 5840533  
; GENERAL INFORMATION:  
; APPLICANT: NIWA, MINEO  
; APPLICANT: SAITO, YOSHIMASA  
; APPLICANT: SASAKI, HITOSHI  
; APPLICANT: HAYASHI, MASAKO  
; APPLICANT: NOTANI, JOUJI  
; APPLICANT: KOBAYASHI, MASAKAZU  
; TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR  
; NUMBER OF SEQUENCES: 67  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
; ADDRESSEE: P.C.  
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400  
; CITY: ARLINGTON  
; STATE: VA  
; COUNTRY: USA  
; ZIP: 22202  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA: US/08/811,949  
; APPLICATION NUMBER: US/08/811,949  
; FILING DATE: 05-MAR-1997  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: OBLON, NORMAN F.  
; REGISTRATION NUMBER: 24,618  
; REFERENCE/DOCKET NUMBER: 18-966-0  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 703-413-3000  
; TELEFAX: 703-413-2220  
; INFORMATION FOR SEQ ID NO: 48:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1314 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: circular  
; MOLECULE TYPE: DNA (genomic)  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 1..1311  
; US-08-811-949-48

Query Match 99.8%; Score 1063.4; DB 2; Length 1314;  
Best Local Similarity 99.9%; Pred. No. 4.2e-291;  
Matches 1064; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 TCTGAGGAAACAGTGACTGCTTCTGGAATGGTCCAGCTACCGTGGCAGCAGCAGC 60

Db 250 TCTGAGGAAACAGTGACTGCTACTTTGGGAATGGGTGAGCTACCGTGGCAGCGACAGC 309  
Qy 61 CTCACCGAGTCGGGTGCTCTCTGCTCCCGTGGAAATTCATGATCCTCATAGGCAAGGTT 120  
Db 310 CTCACCGAGTCGGGTGCTCTCTGCTCCCGTGGAAATTCATGATCCTCATAGGCAAGGTT 369  
Qy 121 TACACAGACAGAACCCAGTCCCGACGAGCTGGGCTGGGCAAAACATAATTAATCTGCGCG 180  
Db 370 TACACAGACAGAACCCAGTCCCGACGAGCTGGGCTGGGCAAAACATAATTAATCTGCGCG 429  
Qy 181 AATCTGTATGGGATGCCAAGCCCTGCTGTCGACAGCTGCTGAAGAACCGCAGGCTGACGTGG 240  
Db 430 AATCTGTATGGGATGCCAAGCCCTGCTGTCGACAGCTGCTGAAGAACCGCAGGCTGACGTGG 489  
Qy 241 GAGTACTGTGATGTCCTCTCTGCTCCACCTGCGGCTGAGACAGTACAGCCAGCTCAG 300  
Db 490 GAGTACTGTGATGTCCTCTCTGCTCCACCTGCGGCTGAGACAGTACAGCCAGCTCAG 549  
Qy 301 TTTCGCATCAAGAGGGGCTTTGCGCGACATGCGCTCCACCCCTGGAGGGCTGCCATC 360  
Db 550 TTTCGCATCAAGAGGGGCTTTGCGCGACATGCGCTCCACCCCTGGAGGGCTGCCATC 609  
Qy 361 TTTCGCATCAAGAGGGGCTTTGCGCGACATGCGCTCCACCCCTGGAGGGCTGCCATC 420  
Db 610 TTTCGCATCAAGAGGGGCTTTGCGCGACATGCGCTCCACCCCTGGAGGGCTGCCATC 669  
Qy 421 TCCTGCTGATGTCCTCTCTGCGGCCACCTGCTTCCAGGAGAGGTTTCCGCCCCACCACTG 480  
Db 670 TCCTGCTGATGTCCTCTCTGCGGCCACCTGCTTCCAGGAGAGGTTTCCGCCCCACCACTG 729  
Qy 481 ACGGTGATCTTGGGAGAACATACCGGTGCTTCCGCGAGGAGGAGGAGGAGGAGGAGGAGG 540  
Db 730 ACGGTGATCTTGGGAGAACATACCGGTGCTTCCGCGAGGAGGAGGAGGAGGAGGAGGAGG 789  
Qy 541 GTCGAAATATACATGTCATTAAGAAATTCGATGATGACACTTACGACAAATGACATTCGG 600  
Db 790 GTCGAAATATACATGTCATTAAGAAATTCGATGATGACACTTACGACAAATGACATTCGG 849  
Qy 601 CTGCTGACGTGAATTCGGATTCGTCGCGCTGTCGCCAGGAGAGGAGGAGGAGGAGGAGGAGG 660  
Db 850 CTGCTGACGTGAATTCGGATTCGTCGCGCTGTCGCCAGGAGAGGAGGAGGAGGAGGAGGAGG 909  
Qy 661 GTGTGCTTCCCGCGGAGACCTGAGCTGCGGAGCTGGAAGGAGGAGGAGGAGGAGGAGGAGG 720  
Db 910 GTGTGCTTCCCGCGGAGACCTGAGCTGCGGAGCTGGAAGGAGGAGGAGGAGGAGGAGGAGG 969  
Qy 721 TACGCAAGCATGAGGCTTGTCTCTTCTATTCGAGAGGCTGGAAGGAGGAGGAGGAGGAGGAGG 780  
Db 970 TACGCAAGCATGAGGCTTGTCTCTTCTATTCGAGAGGCTGGAAGGAGGAGGAGGAGGAGGAGG 1029  
Qy 781 AGACTGTATCCCATCCAGCGCTGACATCAACAATTTACTTTAAACAGAAAGTCAACCGAC 840  
Db 1030 AGACTGTATCCCATCCAGCGCTGACATCAACAATTTACTTTAAACAGAAAGTCAACCGAC 1089  
Qy 841 AACATGCTGTGTGAGAGACATCGGAGCGCGCGGCCCCAGGCAAACTTGACAGAGGCC 900  
Db 1090 AACATGCTGTGTGAGAGACATCGGAGCGCGCGGCCCCAGGCAAACTTGACAGAGGCC 1149  
Qy 901 TGCCAGGCGGATTCGGGAGGCGCCCTGTGTGTCTGAACGATGCGCGGATGACATTTGGTG 960  
Db 1150 TGCCAGGCGGATTCGGGAGGCGCCCTGTGTGTCTGAACGATGCGCGGATGACATTTGGTG 1209  
Qy 961 GGCATCATCAGTGGGCGCTGGGCTGTGACAGAGGAGTGTCCGGGCTGTGACACAAAG 1020  
Db 1210 GGCATCATCAGTGGGCGCTGGGCTGTGACAGAGGAGTGTCCGGGCTGTGACACAAAG 1269  
Qy 1021 GTTACCAACTACCTAGACTGATTCGTGACAACTGACGACCGTGA 1065  
Db 1270 GTTACCAACTACCTAGACTGATTCGTGACAACTGACGACCGTGA 1314

5200340-1  
; Patent No. 5200340  
; APPLICANT: FOSTER, DONALD C.; MULVIHILL, EILEEN R.; O'HARA,  
; PATRICK J.; PINGEL, KURT; YOSHITAKE, SHINJI  
; TITLE OF INVENTION: THROMBIN-ACTIVATED TISSUE PLASMINOGEN  
; ACTIVATORS  
; NUMBER OF SEQUENCES: 34  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/53,412  
; FILING DATE: 22-MAY-1987  
; SEQ ID NO:1:  
; LENGTH: 1738  
5200340-1

Query Match 99.8%; Score 1063.4; DB 6; Length 1738;  
Best Local Similarity 99.9%; Pred. No. 4.8e-291;  
Matches 1064; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 TCTGAGGAAACAGTGACTGCTACTTTGGGAATGGGTGAGCTACCGTGGCAGCGACAGC 60  
Db 640 TCTGAGGAAACAGTGACTGCTACTTTGGGAATGGGTGAGCTACCGTGGCAGCGACAGC 699  
Qy 61 CTCACCGAGTCGGGTGCTCTCTGCTCCCGTGGAAATTCATGATCCTCATAGGCAAGGTT 120  
Db 700 CTCACCGAGTCGGGTGCTCTCTGCTCCCGTGGAAATTCATGATCCTCATAGGCAAGGTT 759  
Qy 121 TACACAGACAGAACCCAGTCCCGACGAGCTGGGCTGGGCAAAACATAATTAATCTGCGCG 180  
Db 760 TACACAGACAGAACCCAGTCCCGACGAGCTGGGCTGGGCAAAACATAATTAATCTGCGCG 819  
Qy 181 AATCTGTATGGGATGCCAAGCCCTGCTGTCGACAGCTGCTGAAGAACCGCAGGCTGACGTGG 240  
Db 820 AATCTGTATGGGATGCCAAGCCCTGCTGTCGACAGCTGCTGAAGAACCGCAGGCTGACGTGG 879  
Qy 241 GAGTACTGTGATGTCCTCTCTGCTCCACCTGCGGCTGAGACAGTACAGCCAGCTCAG 300  
Db 880 GAGTACTGTGATGTCCTCTCTGCTCCACCTGCGGCTGAGACAGTACAGCCAGCTCAG 939  
Qy 301 TTTCGCATCAAGAGGGGCTTTGCGCGACATCGCTTCCACCCCTGGCAGGCTGCCATC 360  
Db 940 TTTCGCATCAAGAGGGGCTTTGCGCGACATCGCTTCCACCCCTGGCAGGCTGCCATC 999  
Qy 361 TTTCGCATCAAGAGGGGCTTTGCGCGACATCGCTTCCACCCCTGGCAGGCTGCCATC 420  
Db 1000 TTTCGCATCAAGAGGGGCTTTGCGCGACATCGCTTCCACCCCTGGCAGGCTGCCATC 1059  
Qy 421 TCCTGCTGATGTCCTCTCTGCGGCCACCTGCTTCCAGGAGGAGGTTTCCGCCCCACCACTG 480  
Db 1060 TCCTGCTGATGTCCTCTCTGCGGCCACCTGCTTCCAGGAGGAGGTTTCCGCCCCACCACTG 1119  
Qy 481 ACGGTGATCTTGGGAGAACATACCGGTGCTTCCCTGCGAGGAGGAGGAGGAGGAGGAGGAGG 540  
Db 1120 ACGGTGATCTTGGGAGAACATACCGGTGCTTCCCTGCGAGGAGGAGGAGGAGGAGGAGGAGG 1179  
Qy 541 GTCGAAATATACATGTCATTAAGAAATTCGATGATGACACTTACGACAAATGACATTCGG 600  
Db 1180 GTCGAAATATACATGTCATTAAGAAATTCGATGATGACACTTACGACAAATGACATTCGG 1239  
Qy 601 CTGCTGACGTGAATTCGGATTCGTCGCGCTGTCGCCAGGAGAGGAGGAGGAGGAGGAGGAGG 660  
Db 1240 CTGCTGACGTGAATTCGGATTCGTCGCGCTGTCGCCAGGAGAGGAGGAGGAGGAGGAGGAGG 1299  
Qy 661 GTGTGCTTCCCGCGGAGACCTGAGCTGCGGAGCTGGAAGGAGTGTGAGCTCTCCGGC 720  
Db 1300 GTGTGCTTCCCGCGGAGACCTGAGCTGCGGAGCTGGAAGGAGTGTGAGCTCTCCGGC 1359  
Qy 721 TACGCAAGCATGAGGCTTGTGTCTCTTCTATTCGAGAGGCGGCTGAAGGAGGCTCATGTC 780  
Db 1360 TACGCAAGCATGAGGCTTGTGTCTCTTCTATTCGAGAGGCGGCTGAAGGAGGCTCATGTC 1419  
Qy 781 AGACTGTATCCCATCCAGCGCTGACATCAACAATTTACTTTAAACAGAAAGTCAACCGAC 840  
Db 1420 AGACTGTATCCCATCCAGCGCTGACATCAACAATTTACTTTAAACAGAAAGTCAACCGAC 1479



QY 841 AACATGCTGTGTCTGGAGACACTCGAGCGCGCGGCCCGAGCAAACTTGCACGACGCC 900  
DB 1480 AACATGCTGTGTCTGGAGACACTCGAGCGCGCGGCCCGAGCAAACTTGCACGACGCC 1539  
QY 901 TGCCAGGCGGATTCGGAGGCGCCCTGTGTGTCTGAACGATGGCCGATGATTTTGTG 960  
DB 1540 TGCCAGGCGGATTCGGAGGCGCCCTGTGTGTCTGAACGATGGCCGATGATTTTGTG 1599  
QY 961 GGCATCATCAGCTGGGCGCTGGGCTGTGGACAGAGAGATGTCCCGGTGTGTACACAAAG 1020  
DB 1600 GGCATCATCAGCTGGGCGCTGGGCTGTGGACAGAGAGATGTCCCGGTGTGTACACAAAG 1659  
QY 1021 GTTACCAACTACTAGATGGATTCGTGACAAACATGCGACCGTGA 1065  
DB 1660 GTTACCAACTACTAGATGGATTCGTGACAAACATGCGACCGTGA 1704

RESULT 11  
US-08-883-795A-39  
; Sequence 39, Application US/08883795A  
; Patent No. 5985607  
; GENERAL INFORMATION:  
; APPLICANT: Delcuve, Genevieve  
; APPLICANT: Awang, Gregor  
; TITLE OF INVENTION: Recombinant DNA Molecules and Expression  
; TITLE OF INVENTION: Vectors for Tissue Plasminogen Activator  
; NUMBER OF SEQUENCES: 39  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: BERESKIN & PARR  
; STREET: 40 King Street West  
; CITY: Toronto  
; STATE: Ontario  
; COUNTRY: Canada  
; ZIP: M5H 3Y2

; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/883,795A  
; FILING DATE: 27-JUN-1997  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Gravelle, Micheline  
; REGISTRATION NUMBER: 40,261  
; REFERENCE/DOCKET NUMBER: 7841-062  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (416) 364-7311  
; TELEFAX: (416) 361-1398  
; INFORMATION FOR SEQ ID NO: 39:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1955 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; ORIGINAL SOURCE:  
; ORGANISM: Homo sapiens  
US-08-883-795A-39

Query Match 99.8%; Score 1063.4; DB 2; Length 1955;  
Best Local Similarity 99.9%; Pred. No. 5e-291;  
Matches 1064; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TCTGAGGAAACAGTGAAGTCTTCTGGGATGGGTGAGCTACCGTGGCAACGACAGC 60  
DB 715 TCTGAGGAAACAGTGAAGTCTTCTGGGATGGGTGAGCTACCGTGGCAACGACAGC 774  
QY 61 CTCACCGAGTGGGGTCTCTGCGCTCCCGTGGGAATTCATGATCCTGATAGGCAAGTT 120  
DB 775 CTCACCGAGTGGGGTCTCTGCGCTCCCGTGGGAATTCATGATCCTGATAGGCAAGTT 834

QY 121 TACACAGCACAGAACCCAGTGCCAGGCACTGGGCTTGGGCAACATAATTTACTGCCGG 180  
DB 835 TACACAGCACAGAACCCAGTGCCAGGCACTGGGCTTGGGCAACATAATTTACTGCCGG 894  
QY 181 AATCCTGATGGGATGCAAGCCCTGTGTGCACTGTGTGAAGAACCGCAGGCTGACGTGG 240  
DB 895 AATCCTGATGGGATGCAAGCCCTGTGTGCACTGTGTGAAGAACCGCAGGCTGACGTGG 954  
QY 241 GAGTACTGTGATGTGCTCTCTGCTCACCCTGGCGGCTGTGACAGACAGTACAGCAGGCTCAG 300  
DB 955 GAGTACTGTGATGTGCTCTCTGCTCACCCTGGCGGCTGTGACAGACAGTACAGCAGGCTCAG 1014  
QY 301 TTTTCGATCAAGAGAGGCTCTTTCGCGACATCCCTCCACCCCTGGCAGGCTGCCATC 360  
DB 1015 TTTTCGATCAAGAGAGGCTCTTTCGCGACATCCCTCCACCCCTGGCAGGCTGCCATC 1074  
QY 361 TTTTCCCAAGCACAGGAGGTGCGCCGAGAGCGGTTCTGTGCGGGGGCATACTCATCAGC 420  
DB 1075 TTTTCCCAAGCACAGGAGGTGCGCCGAGAGCGGTTCTGTGCGGGGGCATACTCATCAGC 1134  
QY 421 TCCTGCTGGATTTCTCTGCGGCGGCTGCTTCCAGAGAGGTTTCGCGGCGGCGGCGGCTG 480  
DB 1135 TCCTGCTGGATTTCTCTGCGGCGGCTGCTTCCAGAGAGGTTTCGCGGCGGCGGCGGCTG 1194  
QY 481 ACGTGTGATCTTGGGCAAGACATACCGGCTGCTCCCTGGCGAGGAGGAGCAGAAATTTGAA 540  
DB 1195 ACGTGTGATCTTGGGCAAGACATACCGGCTGCTCCCTGGCGAGGAGGAGCAGAAATTTGAA 1254  
QY 541 GTCGAAAAATACATTTGTCATAAGGAATTCGATGATGACACTTACGAAATGACATTCGCG 600  
DB 1255 GTCGAAAAATACATTTGTCATAAGGAATTCGATGATGACACTTACGAAATGACATTCGCG 1314  
QY 601 CTGCTGACGTGAATTCGGATTTGTCGCGCTGTGCTCCCGAGAGAGCAGGCTGCTCGCACT 660  
DB 1315 CTGCTGACGTGAATTCGGATTTGTCGCGCTGTGCTCCCGAGAGAGCAGGCTGCTCGCACT 1374  
QY 661 GTGTGCTTCCCGCGGCGGACCTGACGCTGCGGACCTGGACGAGTGTGAGCTCTCCGGC 720  
DB 1375 GTGTGCTTCCCGCGGCGGACCTGACGCTGCGGACCTGGACGAGTGTGAGCTCTCCGGC 1434  
QY 721 TACGCAAGCATGAGGCTTGTCTCTTTCTATTCGAGCGGCTGAAAGGAGGCTCATGTC 780  
DB 1435 TACGCAAGCATGAGGCTTGTCTCTTTCTATTCGAGCGGCTGAAAGGAGGCTCATGTC 1494  
QY 781 AGACTGTACCCATCCAGCGCTGCATCAACATTTACTTACAGACAGTACCGAC 840  
DB 1495 AGACTGTACCCATCCAGCGCTGCATCAACATTTACTTAAACAGACAGTACCGAC 1554  
QY 841 AACATGCTGTGTCTGGAGACACTCGAGGCGCGGCGGCCCGAGGCAAACTTGCACGACGCC 900  
DB 1555 AACATGCTGTGTCTGGAGACACTCGAGGCGCGGCGGCCCGAGGCAAACTTGCACGACGCC 1614  
QY 901 TGCCAGGCGGATTCGGGAGGCGCCCTGCTGTGTCTGAACGATGGCCGCGATGATTTGTG 960  
DB 1615 TGCCAGGCGGATTCGGGAGGCGCCCTGCTGTGTCTGAACGATGGCCGCGATGATTTGTG 1674  
QY 961 GGCATCATCAGCTGGGCGCTGGGCTGTGGACAGAGGATGTCCCGGTGTGTACACAAAG 1020  
DB 1675 GGCATCATCAGCTGGGCGCTGGGCTGTGGACAGAGGATGTCCCGGTGTGTACACAAAG 1734  
QY 1021 GTTACCAACTACTAGATGGATTCGTGACAAACATGCGACCGTGA 1065  
DB 1735 GTTACCAACTACTAGATGGATTCGTGACAAACATGCGACCGTGA 1779

RESULT 12  
5344773-1  
; Patent No. 5344773  
; APPLICANT: WEI, CHA-MER; HSIUNG, NANCY; REDDY, VERNURI B.;  
; LEMONT, JEFFREY F.; DACKOWSKI, WILLIAM; DOUGLAS, RICHARD;  
; COLE, EDWARD S.; PURCELL JR., RICHARD D.; LAU, DAVID TAI-YUI  
; TITLE OF INVENTION: HUMAN UTERINE TISSUE PLASMINOGEN



ACTIVATOR PRODUCED BY RECOMBIANT DNA

NUMBER OF SEQUENCES: 6  
CURRENT APPLICATION DATA:  
FILING DATE: 01-OCT-1985  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 656,770  
FILING DATE: 01-OCT-1984  
SEQ ID NO: 1  
LENGTH: 2457  
5344773-1

Query Match 99.8%; Score 1063.4; DB 6; Length 2457;  
Best Local Similarity 99.9%; Pred. No. 5.5e-291;  
Matches 1064; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 TCTGAGGAAACAGTACTGCTACTTTGGGAATGGTTCAGCTACCGTGGCCACGACAGC 60  
638 TCTGAGGAAACAGTACTGCTACTTTGGGAATGGTTCAGCTACCGTGGCCACGACAGC 697  
61 CTCACCGAGTGGGTGCTCTGCTCCCTCCGCTGGGAATTCATGATCCTGATAGGCAAGTT 120  
698 CTCACCGAGTGGGTGCTCTGCTCCCTCCGCTGGGAATTCATGATCCTGATAGGCAAGTT 757  
121 TACACAGACAGAACCCAGTCCGAGCACTGGGCTGGGCAACATTAATTAATCTGCCGG 180  
758 TACACAGACAGAACCCAGTCCGAGCACTGGGCTGGGCAACATTAATTAATCTGCCGG 817  
181 AATCTGATGGGATGCCAACCCCTGGTGGCCAGCTGCTGAGAACCCGAGCTGACCTGG 240  
818 AATCTGATGGGATGCCAACCCCTGGTGGCCAGCTGCTGAGAACCCGAGCTGACCTGG 877  
241 GAGTACTGTGATGCTCTGCTCCACCTGGCCCTGAGACAGTACAGCCAGCTCAG 300  
878 GAGTACTGTGATGCTCTGCTCCACCTGGCCCTGAGACAGTACAGCCAGCTCAG 937  
301 TTTGCTCAAGAGAGGCTCTTGGCGACATGCTCCACCTGAGGCTGAGGCTGCCATC 360  
938 TTTGCTCAAGAGAGGCTCTTGGCGACATGCTCCACCTGAGGCTGAGGCTGCCATC 997  
361 TTTGCGACAGAGAGGCTGCTGGCGAGAGGCTGCTGCTGGGGGGGATCTATCAGC 420  
998 TTTGCGACAGAGAGGCTGCTGGCGAGAGGCTGCTGCTGGGGGGGATCTATCAGC 1057  
421 TCCTGCTGATGCTCTGCTGGCGGACATGCTTCCAGGAGAGGTTTCCGCCCCACACCTG 480  
1058 TCCTGCTGATGCTCTGCTGGCGGACATGCTTCCAGGAGAGGTTTCCGCCCCACACCTG 1117  
481 ACGTGTATCTTTGGGAGAACATACCGGGTGTCTTCCGAGGAGGAGACAGAAATTTGAA 540  
1118 ACGTGTATCTTTGGGAGAACATACCGGGTGTCTTCCGAGGAGGAGACAGAAATTTGAA 1177  
541 GTCGAATAATATGTCATAGGAATTCATGATGACATGACATGACATGACATGACG 600  
1178 GTCGAATAATATGTCATAGGAATTCATGATGACATGACATGACATGACATGACG 1237  
601 CTGCTGACGCTGAAATCGATTCGCTCCGCTGTGCTCCAGGAGAGCAGCGTGGTCCGACT 660  
1238 CTGCTGACGCTGAAATCGATTCGCTCCGCTGTGCTCCAGGAGAGCAGCGTGGTCCGACT 1297  
661 GTGTGCTCTCCCGGCGGACCTGAGCTGCGGAGCTGAGAGGAGTGTGAGCTCTCCGGC 720  
1298 GTGTGCTCTCCCGGCGGACCTGAGCTGCGGAGCTGAGAGGAGTGTGAGCTCTCCGGC 1357  
721 TACGCGACGATGAGGCTCTGCTCTCTTCTATTCGAGCGGCTGAGAGGCTCATGTC 780  
1358 TACGCGACGATGAGGCTCTGCTCTCTTCTATTCGAGCGGCTGAGAGGCTCATGTC 1417  
781 AGACTGTACCCATCCAGCGCTGACATCAACAATTTTACTTAAACAGAAAGTCCAGC 840  
1418 AGACTGTACCCATCCAGCGCTGACATCAACAATTTTACTTAAACAGAAAGTCCAGC 1477  
841 AACATGCTGTGTGGAGACACTCGGAGCGGCGGCGGCCAGGCAAACTTTCACGACGCC 900

Db 1478 AACATGCTGTGTGTGAGAGACTCGGAGCGCGCGGCCCAAGGAAACTTTCACGACGCC 1537  
Qy 901 TGCCAGGCGGATTCGGGAGGCGCCCTGCTGTGTCTGAACGATGCGCGCATGCTTTGGTG 960  
Db 1538 TGCCAGGCGGATTCGGGAGGCGCCCTGCTGTGTCTGAACGATGCGCGCATGCTTTGGTG 1597  
Qy 961 GGCATCATCAGCTGGGCGCTGGGCTGTGGACAGAGGATGTCCCGGTTGTATACACAAG 1020  
Db 1598 GGCATCATCAGCTGGGCGCTGGGCTGTGGACAGAGGATGTCCCGGTTGTATACACAAG 1657  
Qy 1021 GTTACCAACTACCTAGACTGATTCGTGACACATGCGACCGTGA 1065  
Db 1658 GTTACCAACTACCTAGACTGATTCGTGACACATGCGACCGTGA 1702

RESULT 13

US-08-286-740-1  
Sequence 1, Application US/08286740  
Patent No. 5561053  
GENERAL INFORMATION:  
APPLICANT: Crowley, Craig W.  
TITLE OF INVENTION: METHOD FOR SELECTING HIGH-EXPRESSING  
TITLE OF INVENTION: HOST CELLS  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genentech, Inc.  
STREET: 460 Point San Bruno Blvd  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: patin (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/286,740  
FILING DATE: 05-AUG-1994  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Lee, Wendy M.  
REGISTRATION NUMBER:  
REFERENCE/DOCKET NUMBER: 798  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415/225-1994  
TELEFAX: 415/952-9881  
TELEX: 910/371-7168  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 7360 bases  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
US-08-286-740-1

Query Match 99.8%; Score 1063.4; DB 1; Length 7360;  
Best Local Similarity 99.9%; Pred. No. 8.7e-291;  
Matches 1064; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TCTGAGGAAACAGTACTGCTACTTTGGGAATGGTTCAGCTACCGTGGCCACGACAGC 60  
Db 2246 TCTGAGGAAACAGTACTGCTACTTTGGGAATGGTTCAGCTACCGTGGCCACGACAGC 2305  
Qy 61 CTCACCGAGTGGGTGCTCTGCTCCCTGCGAATTCATGATCCTGATAGGCAAGTT 120  
Db 2306 CTCACCGAGTGGGTGCTCTGCTCCCTGCGAATTCATGATCCTGATAGGCAAGTT 2365  
Qy 121 TACACAGACAGAACCCAGTGGCGCTGGGCAACATAATTAATTAATCTGCGCG 180

```
Db 2366 TACACAGACAGAACCCAGTGGCCAGCAGCTGGSCCTGGCAACATAATTAATCTGCGG 2425
Qy 181 AATCTGTATGGGATGCCAGCCCTGGTGGCCAGTGTCTGAAGAACCGCAGGCTGACGTGG 240
Db 2426 AATCTGTATGGGATGCCAGCCCTGGTGGCCAGTGTCTGAAGAACCGCAGGCTGACGTGG 2485
Qy 241 GAGTACTGTATGTGSCCTCTCTGTCTCCACCTGGSCCTGGACAGTACAGCAGCCTCAG 300
Db 2486 GAGTACTGTATGTGSCCTCTCTGTCTCCACCTGGSCCTGGACAGTACAGCAGCCTCAG 2545
Qy 301 TTTGCGATCAAGAGGAGGCTCTTTCGCGACATCGCTCCACCCCTGGCAGGCTGCCATC 360
Db 2546 TTTGCGATCAAGAGGAGGCTCTTTCGCGACATCGCTCCACCCCTGGCAGGCTGCCATC 2605
Qy 361 TTTGCGACAGCAGAGGAGGCTGGCCGAGAGGCTTCTGTGCGGGGGATATCTATCAGC 420
Db 2606 TTTGCGACAGCAGAGGAGGCTGGCCGAGAGGCTTCTGTGCGGGGGATATCTATCAGC 2665
Qy 421 TCCTGCTGATTTCTCTGCGGCCACATCGCTTTCAGAGAGGCTTTCGCGCCACACCTG 480
Db 2666 TCCTGCTGATTTCTCTGCGGCCACATCGCTTTCAGAGAGGCTTTCGCGCCACACCTG 2725
Qy 481 ACGGTGATCTTTGGCAGAACATACCGGCTGTCTTCGCGAGGAGGAGCAGAAATTTGAA 540
Db 2726 ACGGTGATCTTTGGCAGAACATACCGGCTGTCTTCGCGAGGAGGAGCAGAAATTTGAA 2785
Qy 541 CTCGAAATATATGTCATTAAGAAATTCATGATGACATTCAGCAATGACATTCGC 600
Db 2786 CTCGAAATATATGTCATTAAGAAATTCATGATGACATTCAGCAATGACATTCGC 2845
Qy 601 CTGCTGACGCTGAATCGGATTCGTCGCGCTGTGCGCAGGAGAGCAGCGTGTGCGCACT 660
Db 2846 CTGCTGACGCTGAATCGGATTCGTCGCGCTGTGCGCAGGAGAGCAGCGTGTGCGCACT 2905
Qy 661 GTGTGCTTCCCGCGGACCTGACGCTGCGGAGCTGCGAGAGTGTGAGCTCTCGGC 720
Db 2906 GTGTGCTTCCCGCGGACCTGACGCTGCGGAGCTGCGAGAGTGTGAGCTCTCGGC 2965
Qy 721 TACGCAAGCATAGGCGCTGTCTCTCTTCTATTCGAGCGGCTGAGGAGGCTCATGTC 780
Db 2966 TACGCAAGCATAGGCGCTGTCTCTCTTCTATTCGAGCGGCTGAGGAGGCTCATGTC 3025
Qy 781 AGACTGTACCATCCAGCGCTGACATCAACAATTTACTTAAACAGAAAGTCAACGAC 840
Db 3026 AGACTGTACCATCCAGCGCTGACATCAACAATTTACTTAAACAGAAAGTCAACGAC 3085
Qy 841 AACATGCTGTGTGTGGAGACATCGGAGCGGCGGCCCGGCAAACTTTCAGCAGGCC 900
Db 3086 AACATGCTGTGTGTGGAGACATCGGAGCGGCGGCCCGGCAAACTTTCAGCAGGCC 3145
Qy 901 TGCAGGCGGATTCGGAGGCGCCCTGTGTGTCTGAACGATGCGCGCATGACTTTGGTG 960
Db 3146 TGCAGGCGGATTCGGAGGCGCCCTGTGTGTCTGAACGATGCGCGCATGACTTTGGTG 3205
Qy 961 GGCATCATCAGCTGGGCGTGGCTGTGACAGAGGATGTCGCGGCTGTACACAAAG 1020
Db 3206 GGCATCATCAGCTGGGCGTGGCTGTGACAGAGGATGTCGCGGCTGTACACAAAG 3265
Qy 1021 GTTACCACTACCTAGATGATGATTCGTGACAAACATGCGACCGTGA 1065
Db 3266 GTTACCACTACCTAGATGATGATTCGTGACAAACATGCGACCGTGA 3310
```

## RESULT 14

PCT-US95-09576-1

```
; Sequence 1, Application PC/TUS9509576
; GENERAL INFORMATION:
; APPLICANT: GENENTECH, INC.
; TITLE OF INVENTION: METHOD FOR SELECTING HIGH-EXPRESSING
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
```

```
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: patin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/09576
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/286740
; FILING DATE: 05-AUG-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Lee, Wendy M.
; REGISTRATION NUMBER: 00,000
; REFERENCE/DOCKET NUMBER: 798PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-1994
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7360 bases
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; PCT-US95-09576-1
```

Query Match 99.8%; Score 1063.4; DB 5; Length 7360;

Best Local Similarity 99.9%; Pred. No. 8,7e-291;

Matches 1064; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy 1 TCTGAGGAAACAGTGAATCTTCTGCTTGGGAATGGGTGAGCTTACCTGGCAGCCACAGC 60
Db 2246 TCTGAGGAAACAGTGAATCTTCTGCTTGGGAATGGGTGAGCTTACCTGGCAGCCACAGC 2305
Qy 61 CTACACGAGTGGGTGCTCTGCTCCCTGGAAATTCATGATCTCTGATAGGCAAGGTT 120
Db 2306 CTACACGAGTGGGTGCTCTGCTCCCTGGAAATTCATGATCTCTGATAGGCAAGGTT 2365
Qy 121 TACACAGCAGAACCCAGTGGCCAGGCACTGGGCTGGGCAAAATAATTAATCTGCGG 180
Db 2366 TACACAGCAGAACCCAGTGGCCAGGCACTGGGCTGGGCAAAATAATTAATCTGCGG 2425
Qy 181 AATCTGTATGGGATGCCAGCCCTGTGTCACAGTGTGTCGAGACCGCAGGCTGACGTG 240
Db 2426 AATCTGTATGGGATGCCAGCCCTGTGTCACAGTGTGTCGAGAACCCGAGGCTGACGTG 2485
Qy 241 GAGTACTGTATGTCCTCTCTGCTCCACCTGCGGCTGAGACAGTACAGCAGCCTCAG 300
Db 2486 GAGTACTGTATGTCCTCTCTGCTCCACCTGCGGCTGAGACAGTACAGCAGCCTCAG 2545
Qy 301 TTTGCGATCAAGAGGAGGCTCTTTCGCGACATCGCTCCACCCCTGGCAGGCTGCCATC 360
Db 2546 TTTGCGATCAAGAGGAGGCTCTTTCGCGACATCGCTCCACCCCTGGCAGGCTGCCATC 2605
Qy 361 TTTGCGACAGCAGAGGAGGCTGGCCGAGAGGCTTCTGTGCGGGGGATATCTATCAGC 420
Db 2606 TTTGCGACAGCAGAGGAGGCTGGCCGAGAGGCTTCTGTGCGGGGGATATCTATCAGC 2665
Qy 421 TCCTGCTGATTTCTCTGCGGCCACATCGCTTTCAGAGAGGCTTTCGCGCCACACCTG 480
Db 2666 TCCTGCTGATTTCTCTGCGGCCACATCGCTTTCAGAGAGGCTTTCGCGCCACACCTG 2725
Qy 481 ACGGTGATCTTTGGCAGAACATACCGGCTGTCTTCGCGAGGAGGAGCAGAAATTTGAA 540
Db 2726 ACGGTGATCTTTGGCAGAACATACCGGCTGTCTTCGCGAGGAGGAGCAGAAATTTGAA 2785
```



Qy	841	AACATGCTGTGTCTGGAGACACTCGGAGCGGCGGCCCCAGGCAAACTTGACGACGCC	900
Db	844	AACATGCTGTGTCTGGAGACACTCGGAGCGGCGGCCCCAGGCAAACTTGACGACGCC	903
Qy	901	TGCCAGGCGGATTTCGGAGGCCCCCTGTGTGTCTGAAAGATGGCCGCATGACTTTGGTG	960
Db	904	TGCCAGGCGGATTTCGGAGGCCCCCTGTGTGTCTGAAAGATGGCCGCATGACTTTGGTG	963
Qy	961	GGCATCATCAGCTGGGGCTGGGCTGTGGACAGAGGATGTCCCGGGTGTGTACACAAAG	1020
Db	964	GGCATCATCAGCTGGGGCTGGGCTGTGGACAGAGGATGTCCCGGGTGTGTACACAAAG	1023
Qy	1021	GTTACCAACTACCTAGACTGGATTTCGTGACACACATGCGACCGTGA	1065
Db	1024	GTTACCAACTACCTAGACTGGATTTCGTGACACACATGCGACCGTGA	1068

Search completed: August 19, 2003, 08:35:09  
Job time : 84.3499 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 19, 2003, 04:46:44 ; Search time 300.529 Seconds  
(without alignments)  
7930.701 Million cell updates/sec

Title: US-09-987-455-4  
Perfect score: 1065  
Sequence: 1 tctgaggaacacgtgactg.....gtgacaacatcgaccgtga 1065

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1504479 seqs, 1118970152 residues  
Total number of hits satisfying chosen parameters: 3008958

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications NA.\*  
1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq.\*  
2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq.\*  
3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq.\*  
4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq.\*  
5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq.\*  
6: /cgn2\_6/ptodata/2/pubpna/PCT05\_PUBCOMB.seq.\*  
7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq.\*  
8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq.\*  
9: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq.\*  
10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq.\*  
11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq.\*  
12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq.\*  
13: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq.\*  
14: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq.\*  
15: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq.\*  
16: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq.\*  
17: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1065	100.0	1065	11	US-09-987-455-4
2	1065	100.0	1065	11	US-09-987-455-7
3	1063.4	99.8	1689	9	US-09-969-271-6
4	1063.4	99.8	2509	14	US-10-193-656-7
5	1063.4	99.8	2519	9	US-09-969-271-5
6	1062	99.7	1128	11	US-09-987-455-2
7	1062	99.7	1128	11	US-09-987-455-5
8	1061.8	99.7	2641	10	US-09-974-298-144
9	230.2	21.6	329	12	US-10-007-926A-433
10	191.6	18.0	1212	10	US-09-880-503-15
11	185	17.4	1415	14	US-10-198-846-12748
12	170.8	16.0	1475	9	US-09-735-705-122
13	170.8	16.0	1475	10	US-09-850-716A-122
14	170.8	16.0	1475	10	US-09-897-778-122
15	170.8	16.0	1475	11	US-09-466-396A-122
16	170.8	16.0	1475	12	US-10-117-982-122

17	170.8	16.0	1475	12	US-10-101-510-159	Sequence 159, App
18	170.8	16.0	2336	12	US-09-971-392-46	Sequence 46, Appl
19	170.8	16.0	2336	12	US-10-101-510-458	Sequence 458, App
20	167.6	15.7	1296	13	US-10-076-421-1	Sequence 1, Appli
21	167.6	15.7	2294	9	US-09-735-705-123	Sequence 123, App
22	167.6	15.7	2294	10	US-09-850-716A-123	Sequence 123, App
23	167.6	15.7	2294	10	US-09-897-778-123	Sequence 123, App
24	167.6	15.7	2294	11	US-09-466-396A-123	Sequence 123, App
25	167.6	15.7	2294	12	US-10-117-982-123	Sequence 123, App
26	167.6	15.7	2344	12	US-10-101-510-383	Sequence 383, App
27	166	15.6	474	9	US-09-864-761-1682	Sequence 1682, Ap
28	166	15.6	1236	10	US-09-880-503-12	Sequence 12, Appl
29	166	15.6	2304	12	US-10-301-822-160	Sequence 160, App
30	166	15.6	2304	14	US-10-171-311-183	Sequence 183, App
31	162	15.2	2486	14	US-10-106-698-1989	Sequence 1989, Ap
32	159.8	15.0	831	10	US-09-880-503-14	Sequence 14, Appl
33	159.8	15.0	972	10	US-09-880-503-16	Sequence 16, Appl
34	139	13.1	461	11	US-09-918-995-10405	Sequence 10405, A
35	123	11.5	1959	12	US-10-101-510-743	Sequence 743, App
36	123	11.5	2035	12	US-10-101-510-448	Sequence 448, App
37	123	11.5	2048	11	US-09-858-909-1	Sequence 1, Appli
38	120.4	11.3	2036	10	US-09-954-456-552	Sequence 552, App
39	120.4	11.3	2036	10	US-09-880-107-1612	Sequence 1612, Ap
40	114.4	10.7	116	9	US-09-864-761-18438	Sequence 18438, A
41	105	9.9	457	11	US-09-918-995-11188	Sequence 11188, A
C 42	92.4	8.7	372	10	US-09-920-300A-237	Sequence 237, App
C 43	92.4	8.7	372	13	US-10-033-528-237	Sequence 237, App
44	88	8.3	442	11	US-09-918-995-12715	Sequence 12715, A
45	83.8	7.9	1230	9	US-09-879-792-35	Sequence 35, Appl

ALIGNMENTS

RESULT 1  
US-09-987-455-4  
; Sequence 4, Application US/09987455  
; Publication No. US20030049729A1  
; GENERAL INFORMATION:  
; APPLICANT: Aranya Manosroi  
; APPLICANT: Jiradej Manosroi  
; APPLICANT: Chatchai Tayapiwatana  
; APPLICANT: Friedrich Goetz  
; APPLICANT: Rolf-Guenther Werner  
; TITLE OF INVENTION: Methods for Large Scale Production of Recombinant  
; TITLE OF INVENTION: DNA-Derived tPA or K2S Molecules  
; FILE REFERENCE: 0652.2190001  
; CURRENT APPLICATION NUMBER: US/09/987.455  
; CURRENT FILING DATE: 2001-11-14  
; PRIOR APPLICATION NUMBER: 60/268,574  
; PRIOR FILING DATE: 2001-02-15  
; PRIOR APPLICATION NUMBER: GB 0027779.8  
; PRIOR FILING DATE: 2000-11-14  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 4  
; LENGTH: 1065  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: coding  
; OTHER INFORMATION: sequence for K2S protein  
US-09-987-455-4

Query Match 100.0%; Score 1065; DB 11; Length 1065;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1065; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 TCTCAGGAAACAGTGTCTCTCTTTGGGAATCGGTACCGTGGCCACGACG 60  
Db 1 TCTCAGGAAACAGTGTCTCTCTTTGGGAATCGGTACCGTGGCCACGACG 60  
Qy 61 CTCACCGAGTGGGTGCTCTCTGCTCCGCGGAATTCATGATCTGTAGGCAAGTT 120

```
Db 61 CTCACGAGTGGGTGCTCTCTGCTCCGCTGGAAATTCATGATCCTGATAGGCAAGTT 120
Qy 121 TACACAGACAGAACCCAGTGGCCAGGACCTGGGCTGGCAACATAATTAATCTGCGG 180
Db 121 TACACAGACAGAACCCAGTGGCCAGGACCTGGGCTGGCAACATAATTAATCTGCGG 180
Qy 181 AATCTGATGGGATGCCAAGCCCTGGTGGCCACGCTGCTGAAGAACCGCAGGCTGACGTGG 240
Db 181 AATCTGATGGGATGCCAAGCCCTGGTGGCCACGCTGCTGAAGAACCGCAGGCTGACGTGG 240
Qy 241 GAGTACTGTGATGGCCCTCTCTGCTCCACCTGGCGCTTGAGACAGTACAGCAGCTTCAG 300
Db 241 GAGTACTGTGATGGCCCTCTCTGCTCCACCTGGCGCTTGAGACAGTACAGCAGCTTCAG 300
Qy 301 TTTCGATCAAGAGGAGGCTCTTCGCGACATCGCTCCACCCCTGGCAGGCTGCCATC 360
Db 301 TTTCGATCAAGAGGAGGCTCTTCGCGACATCGCTCCACCCCTGGCAGGCTGCCATC 360
Qy 361 TTTCGACAGCAGGAGGCTCGCCCGAGAGCGGTTCTGTGGGGGGCATACTCATCAGC 420
Db 361 TTTCGACAGCAGGAGGCTCGCCCGAGAGCGGTTCTGTGGGGGGCATACTCATCAGC 420
Qy 421 TCCTGCTGATCTCTCTGCGGCCACTGCTTCCAGGAGAGGTTTCCGCCCCACCCACTG 480
Db 421 TCCTGCTGATCTCTCTGCGGCCACTGCTTCCAGGAGAGGTTTCCGCCCCACCCACTG 480
Qy 481 ACGTGATCTTGGGAGAGACATACCGGTGGTCCCTGCGAGAGGAGGACAGAAATTTGAA 540
Db 481 ACGTGATCTTGGGAGAGACATACCGGTGGTCCCTGCGAGAGGAGGACAGAAATTTGAA 540
Qy 541 GTCGAAAAATACATTTGTCGCGCCCACTGCTTCCAGGAGAGGTTTCCGCCCCACCCACTG 600
Db 541 GTCGAAAAATACATTTGTCGCGCGCCCACTGCTTCCAGGAGAGGTTTCCGCCCCACCCACTG 600
Qy 601 CTGCTGACGCTGAATTCGGATTCGTCCTTCTATTTCGAGCGGCTGAAGAGGCTCATGTC 780
Db 601 CTGCTGACGCTGAATTCGGATTCGTCCTTCTATTTCGAGCGGCTGAAGAGGCTCATGTC 780
Qy 721 TACGCAAGCATGAGGCTTGTCTCTTCTATTTCGAGCGGCTGAAGAGGCTCATGTC 780
Db 721 TACGCAAGCATGAGGCTTGTCTCTTCTATTTCGAGCGGCTGAAGAGGCTCATGTC 780
Qy 781 AGACTGTACCATCCAGCCGCTGCACATCAACAATTTTAAACAGAACAGTCAACGAC 840
Db 781 AGACTGTACCATCCAGCCGCTGCACATCAACAATTTTAAACAGAACAGTCAACGAC 840
Qy 841 AACATGCTGTGTGGAGACACTCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 900
Db 841 AACATGCTGTGTGGAGACACTCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 900
Qy 901 TGCCAGGCGGATTCGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 960
Db 901 TGCCAGGCGGATTCGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 960
Qy 961 GGCATCATCAGTGGGCGCTGGGCTGGAGAGAGGATGTCCCGGGTGTGTACACAAAG 1020
Db 961 GGCATCATCAGTGGGCGCTGGGCTGGAGAGAGGATGTCCCGGGTGTGTACACAAAG 1020
Qy 1021 GTTACCACTACCTAGACTGGATTCTGTGACACATGCGACCGTGA 1065
Db 1021 GTTACCACTACCTAGACTGGATTCTGTGACACATGCGACCGTGA 1065
```

## RESULT 2

US-09-987-455-7

; Sequence 7; Application US/09987455

; Publication No. US20030049729A1

; GENERAL INFORMATION:

```
; APPLICANT: Aranya Manostroi
; APPLICANT: Jiradej Manosroi
; APPLICANT: Chatchai Tavapiwatana
; APPLICANT: Friedrich Goetz
; APPLICANT: Rolf-Guenther Werner
; TITLE OF INVENTION: Methods for Large Scale Production of Recombinant
; FILE REFERENCE: 0652.219001
; CURRENT APPLICATION NUMBER: US/09/987,455
; CURRENT FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: 60/268,574
; PRIOR FILING DATE: 2001-02-15
; PRIOR APPLICATION NUMBER: GB 0027779.8
; PRIOR FILING DATE: 2000-11-14
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 1065
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: coding
; OTHER INFORMATION: sequence for K2S protein
; US-09-987-455-7
```

Query Match 100.0%; Score 1065; DB 11; Length 1065;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1065; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy 1 TCTGAGGAAACAGTGACTGCTACTTTGGAAATGGGTGAGCTACCGTGGCACGACAGC 60
Db 1 TCTGAGGAAACAGTGACTGCTACTTTGGAAATGGGTGAGCTACCGTGGCACGACAGC 60
Qy 61 CTACCCAGTGGGTGCTCTCTGCTCCGCTGGAAATTCATGATCCTGATAGGCAAGTT 120
Db 61 CTACCCAGTGGGTGCTCTCTGCTCCGCTGGAAATTCATGATCCTGATAGGCAAGTT 120
Qy 121 TACACAGACAGAACCCAGTGGCCAGGACCTGGGCTGGCAACATAATTAATCTGCGG 180
Db 121 TACACAGACAGAACCCAGTGGCCAGGACCTGGGCTGGCAACATAATTAATCTGCGG 180
Qy 181 AATCTGATGGGATGCCAAGCCCTGGTGGCCACGCTGCTGAAGAACCGCAGGCTGACGTGG 240
Db 181 AATCTGATGGGATGCCAAGCCCTGGTGGCCACGCTGCTGAAGAACCGCAGGCTGACGTGG 240
Qy 241 GAGTACTGTGATGGCCCTCTCTGCTCCACCTGGCGCTTGAGACAGTACAGCAGCTTCAG 300
Db 241 GAGTACTGTGATGGCCCTCTCTGCTCCACCTGGCGCTTGAGACAGTACAGCAGCTTCAG 300
Qy 301 TTTCGATCAAGAGGAGGCTCTTCGCGACATCGCTCCACCCCTGGCAGGCTGCCATC 360
Db 301 TTTCGATCAAGAGGAGGCTCTTCGCGACATCGCTCCACCCCTGGCAGGCTGCCATC 360
Qy 361 TTTCGACAGCAGGAGGCTCGCCCGAGAGCGGTTCTGTGGGGGGCATACTCATCAGC 420
Db 361 TTTCGACAGCAGGAGGCTCGCCCGAGAGCGGTTCTGTGGGGGGCATACTCATCAGC 420
Qy 421 TCCTGCTGATCTCTCTGCGGCCACTGCTTCCAGGAGAGGTTTCCGCCCCACCCACTG 480
Db 421 TCCTGCTGATCTCTCTGCGGCCACTGCTTCCAGGAGAGGTTTCCGCCCCACCCACTG 480
Qy 481 ACGTGATCTTGGGAGAGACATACCGGTGGTCCCTGCGAGAGGAGGACAGAAATTTGAA 540
Db 481 ACGTGATCTTGGGAGAGACATACCGGTGGTCCCTGCGAGAGGAGGACAGAAATTTGAA 540
Qy 541 GTCGAAAAATACATTTGTCGCGCCCACTGCTTCCAGGAGAGGTTTCCGCCCCACCCACTG 600
Db 541 GTCGAAAAATACATTTGTCGCGCGCCCACTGCTTCCAGGAGAGGTTTCCGCCCCACCCACTG 600
Qy 601 CTGCTGACGCTGAATTCGGATTCGTCCTTCTATTTCGAGCGGCTGAAGAGGCTGCCACT 660
Db 601 CTGCTGACGCTGAATTCGGATTCGTCCTTCTATTTCGAGCGGCTGAAGAGGCTGCCACT 660
Qy 661 GTGTGCTCTCCCGCGGCGGCTGAGCTGCGGAGCTGGAAGAGTGTGAGCTCTCGGC 720
Db 661 GTGTGCTCTCCCGCGGCGGCTGAGCTGCGGAGCTGGAAGAGTGTGAGCTCTCGGC 720
Qy 721 TACGCAAGCATGAGGCTTGTCTCTTCTATTTCGAGCGGCTGAAGAGGCTCATGTC 780
Db 721 TACGCAAGCATGAGGCTTGTCTCTTCTATTTCGAGCGGCTGAAGAGGCTCATGTC 780
Qy 781 AGACTGTACCATCCAGCCGCTGCACATCAACAATTTTAAACAGAACAGTCAACGAC 840
Db 781 AGACTGTACCATCCAGCCGCTGCACATCAACAATTTTAAACAGAACAGTCAACGAC 840
Qy 841 AACATGCTGTGTGGAGACACTCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 900
Db 841 AACATGCTGTGTGGAGACACTCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 900
Qy 901 TGCCAGGCGGATTCGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 960
Db 901 TGCCAGGCGGATTCGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 960
Qy 961 GGCATCATCAGTGGGCGCTGGGCTGGAGAGAGGATGTCCCGGGTGTGTACACAAAG 1020
Db 961 GGCATCATCAGTGGGCGCTGGGCTGGAGAGAGGATGTCCCGGGTGTGTACACAAAG 1020
Qy 1021 GTTACCACTACCTAGACTGGATTCTGTGACACATGCGACCGTGA 1065
Db 1021 GTTACCACTACCTAGACTGGATTCTGTGACACATGCGACCGTGA 1065
```



```
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
; LENGTH: 2509
; TYPE: DNA
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: GenBank / X07393
; DATABASE ENTRY DATE: 1995-03-27
; RELEVANT RESIDUES: (1)...(2509)
US-10-193-656-7

Query Match      99.8%; Score 1063.4; DB 14; Length 2509;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1064; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TCTGAGGAAACAGTACTGCTACTTTGGGAATGGGTGAGCTACCGTGGGACGACAGC 60
DB TCTGAGGAAACAGTACTGCTACTTTGGGAATGGGTGAGCTACCGTGGGACGACAGC 737
QY 61 CTCACCGAGTGGGTGCTCTGCTCCCTCCGCTGGGAATTCATGATCCTGATAGGCAAGTT 120
DB CTCACCGAGTGGGTGCTCTGCTCCCTCCGCTGGGAATTCATGATCCTGATAGGCAAGTT 797
QY 121 TACACAGACAGAAACCCAGTCCCGAGGCACTGGGCTGGGCAACATAATTTACTGCCGG 180
DB TACACAGACAGAAACCCAGTCCCGAGGCACTGGGCTGGGCAACATAATTTACTGCCGG 857
QY 181 AATCTGATGGGATGCCAGCCCTGGTGCACGCTGCTGAAGAACCGCAGGCTGACGTGG 240
DB AATCTGATGGGATGCCAGCCCTGGTGCACGCTGCTGAAGAACCGCAGGCTGACGTGG 917
QY 241 GAGTACTGTGATGCTCCCTCTGCTCCACCTCCGCTGAGACAGTACAGCAGCCTCAG 300
DB GAGTACTGTGATGCTCCCTCTGCTCCACCTCCGCTGAGACAGTACAGCAGCCTCAG 977
QY 301 TTTGCGCATCAAGAGGAGGCTCTTTCGCGACATCGCTCCACCCCTGGCAGGCTGCCATC 360
DB TTTGCGCATCAAGAGGAGGCTCTTTCGCGACATCGCTCCACCCCTGGCAGGCTGCCATC 1037
QY 361 TTTGCGCAGCAGAGAGTCCGCGGAGAGCGGTTCTGTCGGGGGATACATCATCAGC 420
DB TTTGCGCAGCAGAGAGTCCGCGGAGAGCGGTTCTGTCGGGGGATACATCATCAGC 1037
QY 421 TCCTGCTGGATTCTCTGCGCGCCACCTGCTTCAGAGAGAGGTTTCGCGCCCAACACCTG 480
DB TCCTGCTGGATTCTCTGCGCGCCACCTGCTTCAGAGAGAGGTTTCGCGCCCAACACCTG 1157
QY 481 ACGGTGATCTTGGCGAGAACATAACGGGTGGTTCCTGCGGAGGAGGAGCAAAATTTGAA 540
DB ACGGTGATCTTGGCGAGAACATAACGGGTGGTTCCTGCGGAGGAGGAGCAAAATTTGAA 1217
QY 541 GTCGAAATATACATGTCATAGGAATTCGATGATGACATTCAGCAATGACATTCGC 600
DB GTCGAAATATACATGTCATAGGAATTCGATGATGACATTCAGCAATGACATTCGC 1277
QY 601 CTGCTGACGCTGAATTCGATTCGCTGCTGCTCCAGGAGAGCAGCGTGTCCGCACT 660
DB CTGCTGACGCTGAATTCGATTCGCTGCTGCTCCAGGAGAGCAGCGTGTCCGCACT 1337
QY 661 GTGTGCTTCCCGCGGAGACTGTCAGCTGCGGACTGAGAGAGTGTGAGCTCTCCGGC 720
DB GTGTGCTTCCCGCGGAGACTGTCAGCTGCGGACTGAGAGAGTGTGAGCTCTCCGGC 1397
QY 721 TACGCGACGATGAGGCTTGTCTCTTCTATTTCGAGCGGCTGAGAGGCTCATGTC 780
DB TACGCGACGATGAGGCTTGTCTCTTCTATTTCGAGCGGCTGAGAGGCTCATGTC 1457
QY 781 AGACTGTACCATCCAGCGCTGCACATCAACATTTTACTTTAAGAGAGTCAACGAC 840
DB AGACTGTACCATCCAGCGCTGCACATCAACATTTTACTTTAAGAGAGTCAACGAC 1517
QY 841 AACATGCTGTGTGGAGACACTCGGAGCGCGGCGCCAGGCAAACTTTCAGCAGCGC 900
DB AACATGCTGTGTGGAGACACTCGGAGCGCGGCGCCAGGCAAACTTTCAGCAGCGC
```

```
DB 1518 AACATGCTGTGTCTGTGAGACACTCGGAGCGCGGCGCCAGCAAACTTTCAGCAGCGC 1577
QY 901 TGCCAGGCGGATTCGGGAGGCGCCCTGGTGTGTCTGAACGATGCGCGCATGACTTTGGTG 960
DB TGCCAGGCGGATTCGGGAGGCGCCCTGGTGTGTCTGAACGATGCGCGCATGACTTTGGTG 1637
QY 961 GGCATCATCAGCTGGGCGCTGGGCTGTGGACAGAGGATGTCCCGGGTGTGTACACAAAG 1020
DB GGCATCATCAGCTGGGCGCTGGGCTGTGGACAGAGGATGTCCCGGGTGTGTACACAAAG 1697
QY 1021 GTTACCAACTACTAGTGTGATTCGATTCGTGACAAACATGCGACCGTGA 1065
DB GTTACCAACTACTAGTGTGATTCGTGACAAACATGCGACCGTGA 1742

RESULT 5
US-09-969-271-5
; Sequence 5, Application US/09969271
; Patent No. US20020098179A1
; GENERAL INFORMATION:
; APPLICANT: Pfizer Inc. (All designated States except GB and EP (GB));
; APPLICANT: Pfizer Limited (GB and EP (GB) only)
; TITLE OF INVENTION: Pharmaceutical Combinations
; FILE REFERENCE: PCS10951APME
; CURRENT APPLICATION NUMBER: US/09/969,271
; CURRENT FILING DATE: 2001-10-01
; PRIOR APPLICATION NUMBER: GB 0025473.0
; PRIOR FILING DATE: 2000-10-17
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 2519
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-969-271-5

Query Match      99.8%; Score 1063.4; DB 9; Length 2519;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1064; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TCTGAGGAAACAGTACTGCTACTTTGGGAATGGGTGAGCTACCGTGGGACGACAGC 60
DB TCTGAGGAAACAGTACTGCTACTTTGGGAATGGGTGAGCTACCGTGGGACGACAGC 760
QY 61 CTCACCGAGTGGGTGCTCTGCTCCCTCCGCTGGGAATTCATGATCCTGATAGGCAAGTT 120
DB CTCACCGAGTGGGTGCTCTGCTCCCTCCGCTGGGAATTCATGATCCTGATAGGCAAGTT 820
QY 121 TACACAGCAGAACCCAGTCCCGAGGCACTGGGCTGGGCAAAATAATTTACTGCCGG 180
DB TACACAGCAGAACCCAGTCCCGAGGCACTGGGCTGGGCAAAATAATTTACTGCCGG 880
QY 181 AATCTGATGGGATGCCAGCCCTGGTGCACGCTGCTGAAGAAACCGCAGGCTGACGTGG 240
DB AATCTGATGGGATGCCAGCCCTGGTGCACGCTGCTGAAGAAACCGCAGGCTGACGTGG 940
QY 241 GAGTACTGTGATGTCCTCTCTGCTCCACCTGCGGCTGAGACAGTACAGCAGCCTCAG 300
DB GAGTACTGTGATGTCCTCTCTGCTCCACCTGCGGCTGAGACAGTACAGCAGCCTCAG 1000
QY 301 TTTGCGCATCAAGAGGAGGCTCTTTCGCGGACATGCGCTCCACCCCTGGCAGGCTGCCATC 360
DB TTTGCGCATCAAGAGGAGGCTCTTTCGCGGACATGCGCTCCACCCCTGGCAGGCTGCCATC 1060
QY 361 TTTGCGCAGCAGAGGCTGCGCGGAGAGCGGTTCTGTCGGGGGATACATCATCAGC 420
DB TTTGCGCAGCAGAGGCTGCGCGGAGAGCGGTTCTGTCGGGGGATACATCATCAGC 1120
QY 421 TCCTGCTGGATTCTCTGCGCGCCACTGCTTCAGAGAGGTTTCGCGCCCAACACCTG 480
DB TCCTGCTGGATTCTCTGCGCGCCACTGCTTCAGAGAGGTTTCGCGCCCAACACCTG 1180
QY 481 ACGGTGATCTTGGCGAGAACATAACCGGCTGGTTCCTTCGCGGAGGAGGAGCAAAATTTGAA 540
```





## RESULT 7

US-09-987-455-5  
 ; Sequence 5, Application US/09987455  
 ; Publication No. US20030049729A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Aranya Manosroi  
 ; APPLICANT: Jitadej Manosroi  
 ; APPLICANT: Chatchai Tayapiwatana  
 ; APPLICANT: Friedrich Goetz  
 ; APPLICANT: Rolf-Guenther Werner  
 ; TITLE OF INVENTION: Methods for Large Scale Production of Recombinant  
 ; TITLE OF INVENTION: DNA-Derived tPA or K2S Molecules  
 ; FILE REFERENCE: 0652.2190001  
 ; CURRENT APPLICATION NUMBER: US/09/987,455  
 ; CURRENT FILING DATE: 2001-11-14  
 ; PRIOR APPLICATION NUMBER: 60/268,574  
 ; PRIOR FILING DATE: 2001-02-15  
 ; PRIOR APPLICATION NUMBER: GB 0027779.8  
 ; PRIOR FILING DATE: 2000-11-14  
 ; NUMBER OF SEQ ID NOS: 25  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 5  
 ; LENGTH: 1128  
 ; TYPE: DNA  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: Description of Artificial Sequence: coding  
 ; OTHER INFORMATION: sequence for OmpA-K2S fusion protein  
 US-09-987-455-5

Query Match 99.7%; Score 1062; DB 11; Length 1128;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1062; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 TCTGAGGAAACAGTGACTGCTTCTGGGAATGGGTGAGCTACCGTGGCAGCACAGC 60
DB 67 TCTGAGGAAACAGTGACTGCTTCTGGGAATGGGTGAGCTACCGTGGCAGCACAGC 126
QY 61 CTCACCGAGTGGGTGCTCTGCTCCGCTGGGAATTCATGATCCTGATAGGCAAGTT 120
DB 127 CTCACCGAGTGGGTGCTCTGCTCCGCTGGGAATTCATGATCCTGATAGGCAAGTT 186
QY 121 TACACGACAGAACCCAGTGCCAGGACCTGGGCTGGGCAACATAATTAATCTGCGG 180
DB 187 TACACGACAGAACCCAGTGCCAGGACCTGGGCTGGGCAACATAATTAATCTGCGG 246
QY 181 AATCTGTATGGGATGCCAAGCCCTGTGTGCCACGCTGTGAAGAACCGCAGGCTGACGTGG 240
DB 247 AATCTGTATGGGATGCCAAGCCCTGTGTGCCACGCTGTGAAGAACCGCAGGCTGACGTGG 306
QY 241 GAGTACTGTATGTCCTCTCTGCTCCACCTGGCGGCTTGAGACAGTACAGCAGCCTCAG 300
DB 307 GAGTACTGTATGTCCTCTCTGCTCCACCTGGCGGCTTGAGACAGTACAGCAGCCTCAG 366
QY 301 TTTCGCATCAAGAGGAGGCTCTTCGCCACATCGCTCCACCCCTGGCAGGCTGCCATC 360
DB 367 TTTCGCATCAAGAGGAGGCTCTTCGCCACATCGCTCCACCCCTGGCAGGCTGCCATC 426
QY 361 TTTCGCATCAAGAGGAGGCTCTTCGCCACATCGCTCCACCCCTGGCAGGCTGCCATC 420
DB 427 TTTCGCATCAAGAGGAGGCTCTTCGCCACATCGCTCCACCCCTGGCAGGCTGCCATC 486
QY 421 TCCTGCTGGATTCCTCTGCGGCCACATGCTTCAGAGAGAGGTTTCGCCGCCACACCTG 480
DB 487 TCCTGCTGGATTCCTCTGCGGCCACATGCTTCAGAGAGAGGTTTCGCCGCCACACCTG 546
QY 481 ACGGTGATCTTGGGACAGAACATACCGGTGTCTTCCTGGCAGGAGGACAGAAATTTGAA 540
DB 547 ACGGTGATCTTGGGACAGAACATACCGGTGTCTTCCTGGCAGGAGGACAGAAATTTGAA 606
QY 541 CTCGAAAAATACATTTGTCATTAAGGAATTCGATGATGACCTTACGACAAATGACATTCGG 600

```

```

DB 607 GTCGAAAAATACATTTGTCATTAAGGAATTCGATGATGACACTTACGACAAATGACATTCGG 666
QY 601 CTGCTGAGCTGAAATCGGAATTCGTCCTCGCTGTCGCCAGAGAGCAGCGTGGTCCGCACT 660
DB 667 CTGCTGAGCTGAAATCGGAATTCGTCCTCGCTGTCGCCAGAGAGCAGCGTGGTCCGCACT 726
QY 661 GTGTGCTTCTCCCGCGGAGACCTGTCAGCTCCGAGCTGGACGGAGTGTGAGCTCTCCGGC 720
DB 727 GTGTGCTTCTCCCGCGGAGACCTGTCAGCTCCGAGCTGGACGGAGTGTGAGCTCTCCGGC 786
QY 721 TAGCGCAAGCATGAGGCTTGTCTCTTTTATTCGAGCGGCTGAAGAGGCTCATGTC 780
DB 787 TAGCGCAAGCATGAGGCTTGTCTCTTTTATTCGAGCGGCTGAAGAGGCTCATGTC 846
QY 781 AGACTGTACCATCCAGCGCTGCACATCAACATTTACTTAACAGAACAGTACACGAC 840
DB 847 AGACTGTACCATCCAGCGCTGCACATCAACATTTACTTAACAGAACAGTACACGAC 906
QY 841 AACATGCTGTGTGCTGGAGACACTCGGAGCGCGGGCCCCAGGCAAACTTGCACGACGCC 900
DB 907 AACATGCTGTGTGCTGGAGACACTCGGAGCGCGGGCCCCAGGCAAACTTGCACGACGCC 966
QY 901 TGCAGGCGGATTCGGAGAGCCCCCTGTGTGTCTGAACGATGCCCGCATGCTTTGGTG 960
DB 967 TGCAGGCGGATTCGGAGAGCCCCCTGTGTGTCTGAACGATGCCCGCATGCTTTGGTG 1026
QY 961 GGCAATCATGCTGGGCTGGGCTGTGGACAGAGAGATGTCCTGGGTGTGTACACAAAG 1020
DB 1027 GGCAATCATGCTGGGCTGGGCTGTGGACAGAGAGATGTCCTGGGTGTGTACACAAAG 1086
QY 1021 GTTACCAACTACCTAGACTGGATTGCGACAAACATGCGACCG 1062
DB 1087 GTTACCAACTACCTAGACTGGATTGCGACAAACATGCGACCG 1128

```

## RESULT 8

US-09-974-298-144  
 ; Sequence 144, Application US/09974298  
 ; Patent No. US20020156263A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Chen, Huel-Mel  
 ; TITLE OF INVENTION: GENES EXPRESSED IN BREAST CANCER  
 ; FILE REFERENCE: PA-0037 P  
 ; CURRENT APPLICATION NUMBER: US/09/974,298  
 ; CURRENT FILING DATE: 2001-10-04  
 ; PRIOR APPLICATION NUMBER: 60/238,331  
 ; PRIOR FILING DATE: 2000-05-10  
 ; NUMBER OF SEQ ID NOS: 194  
 ; SOFTWARE: PERL Program  
 ; SEQ ID NO 144  
 ; LENGTH: 2641  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: misc feature  
 ; OTHER INFORMATION: Incyte ID No. US20020156263A1 1001470CB1  
 ; NAME/KEY: unsure  
 ; LOCATION: 2635  
 ; OTHER INFORMATION: a, t, c, g, or other  
 US-09-974-298-144

Query Match 99.7%; Score 1061.8; DB 10; Length 2641;  
 Best Local Similarity 99.8%; Pred. No. 0;  
 Matches 1063; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

QY 1 TCTGAGGAAACAGTGACTGCTTCTGGGAATGGGTGAGCTACCGTGGCAGCACAGC 60
DB 821 TCTGAGGAAACAGTGACTGCTTCTGGGAATGGGTGAGCTACCGTGGCAGCACAGC 880
QY 61 CTCACCGAGTGGGTGCTCTGCTCCCGTGGAAATTCATGATCCTGATAGGCAAGTT 120
DB 881 CTCACCGAGTGGGTGCTCTGCTCCCGTGGAAATTCATGATCCTGATAGGCAAGTT 940

```



.....

; GENERAL INFORMATION:

Patent No. US20020052329A1  
GENERAL INFORMATION.

```
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy
; APPLICANT: Fanger, Gary R.
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aljun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Henderson, Robert A.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Fanger, Neil
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C14
; CURRENT APPLICATION NUMBER: US/09/735,705
; CURRENT FILING DATE: 2000-12-12
; NUMBER OF SEQ ID NOS: 419
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 122
; LENGTH: 1475
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-735-705-122

Query Match      16.0%; Score 170.8; DB 9; Length 1475;
Best Local Similarity 51.3%; Pred. No. 7e-44;
Matches 551; Conservative 0; Mismatches 457; Indels 66; Gaps 4;

Qy 18 CTGCTACTTTGGGAATGGTTCAGCTACGCTTGGGCGACGACAGCCCTCACCGAGTCGGGTGC 77
Db |||||
Qy 287 CTGCTATGAGGGAAATGGTTCACCTTTTACCGAGGAAAGCCAGCACCTGACACCATGGGCGC 346
Db |||||
Qy 78 CTCTGCTCCCGTGGAAATTCATGATCTCTGATAGGCAAGTTTACACAGACACAGAACCC 137
Db |||||
Qy 347 GCCCTGCTCCCGTGGAAATTCATGATCTCTGATAGGCAAGTTTACACAGACACAGAACCC 406
Db |||||
Qy 138 CAGTCCCGACGACCTGGGCGTGGGCAAAATAATTTACTCCGGAATCTGTATGGGATGC 197
Db |||||
Qy 407 TGATGCTCTTCAGCTGGGCGTGGGCAAAATAATTTACTGCAGGAACCCAGACCAACCGGAG 466
Db |||||
Qy 198 CAAGCCCTGGTCCAGCTGCTGAAGAACCGAGGCTGACGTTGGGAGTACTGTGATGTGCC 257
Db |||||
Qy 467 GCGACCCCTGGTGTATGTGACAGGTGGGCGCTTAAAGCCGCTTGTCCAGAGTGCATGGTGA 526
Db |||||
Qy 258 CTCTGCTCCACCTCGCGGCTGAG----- 281
Db |||||
Qy 527 TGACTGGCAGATGGAAGAAAGCCCTCTCTCTCCAGAGAAATTAATAATTCAGTGTGG 586
Db |||||
Qy 282 ---ACAGTACAGCCAGCCTCAGTTTCGCATCAAAGGAGGGCTCTTTCGCGGACATCGCCTC 338
Db |||||
Qy 587 CCAAAAGACTCTGAGGCGCGCTTTAAGATTATTGGGGGAGAAATTCACCACCATCGAGAA 646
Db |||||
Qy 339 CCACCCCTGGCAGGTGCGATCTTTCGCAAGCAGAGAGGTGCGCCCGAGAGCGGTTCTT 398
Db |||||
Qy 647 CCAGCCCTGGTTTGGCGGCATCTACAGAGAGGACCGGGG---GGGCTCTGTCACTACGT 703
Db |||||
Qy 399 GTGCGGGGCGATPACTCATCAGCTCTGTGGAATTTCTCTCGCGCCCACTGCTTCCAGGA 458
Db |||||
Qy 704 GTGTGAGGAGCCCTCATCAGCCCTTGTGGTGTATCAGCGCCACACACTGCTTCATTTGA 763
Db |||||
Qy 459 GAGGTTTCCGCCCCACCACTGACGGTGATCTTGGGCGAGAAACATACCGGGTGGTCCCTGG 518
Db |||||
Qy 764 TTATCCAAAAGAGGAGGACTACATGCTCTACCTGGTCTGCTCAAGGCTTAACCTCAACAC 823
Db |||||
Qy 519 CGAGAGAGCAGAAATTTGAAGTCGAAAAATAATTTGTCATTAAGGAATTCGATGATGA 578
Db |||||
Qy 824 GCAAGGGAGATGAAGTTTGGAGTGGAAAAACCTCATCTACACAGGACTACAGCGCTGA 883
Db |||||
Qy 579 CA-----CTTACGACATGATGCGCTGCTGAGCTGAAATCGGATTTGTCGCCGTG 632
Db |||||
Qy 884 CAGCTTGTCTACCAACAGACATTCCTTGTCTGAAGATCCGTTTCCAAAGGAGGCGAGGTG 943
Db |||||

RESULT 13
US-09-850-716A-122
; Sequence 122, Application US/09850716A
; Patent No. US20020115139A1
; GENERAL INFORMATION:
; APPLICANT: Kalos, Michael D.
; APPLICANT: McNeill, Marc W.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C15
; CURRENT APPLICATION NUMBER: US/09/850,716A
; CURRENT FILING DATE: 2001-05-07
; NUMBER OF SEQ ID NOS: 440
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 122
; LENGTH: 1475
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-850-716A-122

Query Match      16.0%; Score 170.8; DB 10; Length 1475;
Best Local Similarity 51.3%; Pred. No. 7e-44;
Matches 551; Conservative 0; Mismatches 457; Indels 66; Gaps 4;

Qy 18 CTGCTACTTTGGGAATGGTTCAGCTACGCTTGGGCGACGACAGCCCTCACCGAGTCGGGTGC 77
Db |||||
Qy 287 CTGCTATGAGGGAAATGGTTCACCTTTTACCGAGGAAAGCCAGCACCTGACACCATGGGCGC 346
Db |||||
Qy 78 CTCTGCTCCCGTGGAAATTCATGATCTCTGATAGGCAAGTTTACACAGACACAGAACCC 137
Db |||||
Qy 347 GCCCTGCTCCCGTGGAAATTCATGATCTCTGATAGGCAAGTTTACACAGACACAGAACCC 406
Db |||||
Qy 138 CAGTCCCGACGACCTGGGCGTGGGCAAAATAATTTACTCCGGAATCTGTATGGGATGC 197
Db |||||
Qy 407 TGATGCTCTTCAGCTGGGCGTGGGCAAAATAATTTACTGCAGGAACCCAGACCAACCGGAG 466
Db |||||
Qy 198 CAAGCCCTGGTCCAGCTGCTGAAGAACCGAGGCTGACGTTGGGAGTACTGTGATGTGCC 257
Db |||||
Qy 467 GCGACCCCTGGTGTATGTGACAGGTGGGCGCTTAAAGCCGCTTGTCCAGAGTGCATGGTGA 526
Db |||||
Qy 258 CTCTGCTCCACCTCGCGGCTGAG----- 281
Db |||||
Qy 527 TGACTGGCAGATGGAAGAAAGCCCTCTCTCTCCAGAGAAATTAATAATTCAGTGTGG 586
Db |||||
```



Db 1181 -----GAAACAGATTCTGCGCAGGAGACTCAGGGGACCCCTCGTCTG 1225  
Qy 933 TCTGAACGATGCCGCGATGCTTTGGTGGGATCATCAGCTGGGCGCTGGGCTGTGGACA 992  
Db 1226 TTCCCTCCAAAGCCGCGATGCTTTGACTGGAATTGTGAGCTGGGCGCGTGGATGTGCCCT 1285  
Qy 993 GAAGGATGTCCGGGTGTGTACAAAGTTACCAACTACCTAGACTCGATTTCG 1046  
Db 1286 GAAGGACAAAGCCAGGCGGTCTACAGAGAGTCTCACACTCTTTACCTCGATCCG 1339

RESULT 15  
US-09-466-396A-122  
; Sequence 122, Application US/09466396A  
; Publication No. US20030119763A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Tonglong  
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY AND  
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER  
; FILE REFERENCE: 210121.455C4  
; CURRENT APPLICATION NUMBER: US/09/466,396A  
; CURRENT FILING DATE: 1999-12-17  
; NUMBER OF SEQ ID NOS: 224  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 122  
; LENGTH: 1475  
; TYPE: DNA  
; ORGANISM: Homo sapien  
US-09-466-396A-122

Query Match 16.0%; Score 170.8; DB 11; Length 1475;  
Best Local Similarity 51.3%; Pred. No. 7e-44;  
Matches 551; Conservative 0; Mismatches 457; Indels 66; Gaps 4;

Qy 18 CTGCTACTTTGGGATGGGTACGCTACGCTGGGCGGACGACGCTCACCGAGTCGGGTGC 77  
Db 287 CTGCTATGAGGGAATGGTCTCACTTTTACCGAGGAAGCGCCAGCACTGACCACTGGGCGG 346  
Qy 78 CTCTGCTCCTCCGCTGGAAATTCATGATCTGTATAGGCAAGGTTTACACAGCACAGAACCC 137  
Db 347 GCCCTGCTCCTCGGAACTGTGCCACTGTCTTCCAGCAAGCTACCATGCCACAGATC 406  
Qy 138 CAGTCCCGCAGGCACTGGGCGCTGGGCAAAACATAATTAATCTGCCGGAATCCTGATGGGGATGC 197  
Db 407 TGATGCTCTTCACTGGGCGCTGGGGAACATAATTAATCTGAGGAGACCCAGACAAACCGGAG 466  
Qy 198 CAAGCCCTGGTCCAGCTGTGTGAAGAACCGGAGGCTGAGCTGGGAGTACTGTGATGTGCC 257  
Db 467 GCGACCTCTGGTGTATGTGAGGTTGGGCTTAAAGCCGCTTGTCCAAAGAGTGTGATGTGCA 526  
Qy 258 CTCCTGCTCCACCTCGCGGCTGTAG----- 281  
Db 527 TGACTGGCGAGATGAAAAAGCCCTCTCTCTCCAGAGAAATTAATAATTCAGTGTGG 586  
Qy 282 ---ACAGTACAGCCAGCTCAGTTTCGATCAAGAGAGGCTCTTCGCCGACATCGCCTC 338  
Db 587 CCMAAAGACTCTGAGGCGCCGCTTTAAGATTATTGGGGAGAAATTCACCAACCATCGAGAA 646  
Qy 339 CCACCCCTGGCAGGCTGCACTTTTGGCCAGACAGAGAGGTGCGCCCGGAGAGCGGTTTCCT 398  
Db 647 CCAGCCCTGGTTTGGGCGCATCTACAGGAGGACCGGGG---GGGCTCTGTCAACCTACGT 703  
Qy 399 GTGCGGGGCGATCTCATCAGCTCTGCTGATTTCTCTCGCGCCCACTGCTTCCAGGA 458  
Db 704 GTGTGAGGCGAGCTCATCAGCCCTTGTGGTGTATCAGCCCAACACTGCTTCAATGA 763  
Qy 459 GAGGTTTCCGCCCCACCTGACCGGTGATCTTTGGGCGAGAACATACCGGGTGGTCCCTGG 518  
Db 764 TTACCCAAAGAGGAGGAGTACATGCTTACCTGGTGTCTCAAGGCTTAACCTCAACAC 823  
Qy 519 CGAGGAGGAGCGAAATTTGAAGTCGAAAAATACATTTGTCATTAAGGAATTCGATGA 578

Db 824 GCAAGGGGAGATGAAGTTTGAAGGTGGAAAAACCTCATCTCTACACAGGACTACAGCGCTGA 883  
Qy 579 CA-----CTTACGACAAATGACATTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 632  
Db 884 CACGCTTGTCTCACCAACAGACATTTGCTTGTGAAGATCCGTTTCCAAGGAGGCGAGGTG 943  
Qy 633 TGCCCGAGAGAGCAGCGTGGTCCGCACTGTGTGCTTCCCGCCCGCGGACCTTGCAGCTGCC 692  
Db 944 TGCAGCGCATCCCGGACTATACAGACCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1003  
Qy 693 GGAAGTGAAGGAGTGTGAGCTCTCCGGCTAGCGCAAGCATGAGGCGCTTGTCTCTCTTCTA 752  
Db 1004 GTTTGGCACAAGCTGTGAGATCACTGGCTTTGGAAAAAGAGAAATTTACCGGACTATCTCTA 1063  
Qy 753 TTCGGAGCGGCTGAAGGAGGCTCATGTGACAGCTGTATCCCATCCAGCCGCTGCACATCACA 812  
Db 1064 TCCGGAGCAGCTGAAGATGACTGTGTGAAGCTGATTTCCACCGGAGTGTCAAGCAGCC 1123  
Qy 813 ACATTTACTTAAAGAAAGACTCAGGACCAAGCATGCTGTGTGCTGTGAGACACTCGAGCGG 872  
Db 1124 CCACCTACTACGGCTCTTGAAGTCCACCAAAATGCTGTGTGCTGTGCTGCTGCTGCTGCTG 1180  
Qy 873 CGGGCCCCCAGGCAAACTTGCACAGCGCTGCCAGGCGGATTCGGGAGGCGCCCTGCTGTG 932  
Db 1181 -----GAAAAACAGATTCTGCGCAGGAGACTCAGGGGAGACCCCTCGTCTG 1225  
Qy 933 TCTGAACGATGGCGCGCATGACTTTTGTGGGCGATCATCAGCTGGGCGCTTGGGCTGTGGACA 992  
Db 1226 TTCCCTCCAAAGCCGCGATGACTTTGACTGGAAATTTGTGAGCTGGGCGCGTGGATGTGCCCT 1285  
Qy 993 GAAGGATGTCCCGGCTGTGTACACAAAGGTTTACCAACTACCTAGACTGGAATTCG 1046  
Db 1286 GAAGGACAAAGCCAGGCGCTTACAGAGAGTCTCACACTTCTTTACCTCGATCCG 1339

Search completed: August 19, 2003, 14:22:25  
Job time : 302.529 secs





GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 18, 2003, 22:44:33 ; Search time 3115.64 Seconds  
(without alignments)  
8307.845 Million cell updates/sec

Title: US-09-987-455-4  
Perfect score: 1065  
Sequence: 1 tctgagggaacagtactg.....gtgacaacatcgaccgtga 1065

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0  
Searched: 22781392 seqs, 12152238056 residues  
Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST:\*  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hic:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_est3:\*  
12: gb\_est4:\*  
13: gb\_est5:\*  
14: gb\_est6:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: em\_gss\_hum:\*  
18: em\_gss\_inv:\*  
19: em\_gss\_pln:\*  
20: em\_gss\_vit:\*  
21: em\_gss\_fun:\*  
22: em\_gss\_mam:\*  
23: em\_gss\_mus:\*  
24: em\_gss\_pro:\*  
25: em\_gss\_rod:\*  
26: em\_gss\_phg:\*  
27: em\_gss\_vrl:\*  
28: gb\_gss1:\*  
29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	895	84.0	930	13	BX329047
2	853	80.1	912	13	BUI49958
3	850.6	79.9	962	13	BX389609
4	818.6	76.9	891	13	BUI46167

5	815.8	76.6	993	13	BQ278177
6	811.2	75.2	883	13	BQ690978
7	807.6	75.8	918	13	BX389608
8	796.2	74.8	930	13	BUI45014
9	782.8	73.5	888	13	BUI57720
10	787.4	72.1	916	12	BI765247
11	785.2	71.8	950	13	BQ278999
12	706.8	66.4	948	10	BE616613
13	697.4	65.5	958	12	BG763582
14	694.4	65.2	974	13	BQ960216
15	693.4	65.1	709	10	BE732704
16	685.2	64.3	1060	12	BM556999
17	677	63.6	1042	10	BE732413
18	672	63.1	957	13	BQ687538
19	668.8	62.8	948	13	BQ687779
20	665.2	62.5	704	9	AUI134301
21	658.2	61.8	903	13	BQ689095
22	656.2	61.6	830	13	BQ896938
23	655.2	61.5	904	13	BQ689840
24	655.2	61.5	911	13	BQ684734
25	654.6	61.5	867	13	BUI79903
26	652.4	61.3	818	9	AUI124602
27	651	61.1	818	12	BM006474
28	649.6	61.0	797	13	BQ879911
29	642.2	60.3	1037	13	BUI57493
30	631.2	59.3	813	10	BG120780
31	629.8	59.1	663	9	AUI134463
32	624	58.6	624	10	BE549489
33	585.4	55.0	587	10	BE394944
34	561	52.7	561	9	AL047284
35	553.4	52.0	555	14	CB128855
36	535.4	50.3	537	14	CB155052
37	531	49.9	964	12	BG769384
38	517.4	48.6	905	13	BUI59271
39	514.8	48.3	883	14	CD359959
40	514.4	48.3	516	14	CB140923
41	496.4	46.6	916	13	BUI57241
42	487.8	45.8	527	9	AW630119
43	485.8	45.6	553	9	AW394127
44	477.8	44.9	917	13	BUI90904
45	475.8	44.7	937	13	BQ679501

ALIGNMENTS

RESULT 1  
BX329047/c  
LOCUS  
DEFINITION BX329047 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA clone CSQD1039YK15 3-PRIME, mRNA sequence.  
ACCESSION BX329047  
VERSION BX329047.1 GI:30307809  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE  
AUTHORS Li, W.B., Gruber, C., Jesses, J. and Polayes, D.  
TITLE Full-length cDNA libraries and normalization  
JOURNAL Unpublished  
COMMENT  
Genoscope - Centre National de Sequencage  
BP 191 91006 EVRY cedex - France  
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 4498.r For more information about this cluster, see  
http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CS0BAI0212D08\_C  
S01980 lcluster=4498.r. Contact : Feng Liang Email :  
fliang@lifetech.com URL : http://fulllength.invitrogen.com/  
Invitrogen Corporation 1600 Faraday Avenue Genoscope sequence ID :

```
CSOBAI0212D08_CS01980.1.
FEATURES
  Location/Qualifiers
    1..930
      /organism="Homo sapiens"
      /mol_type="mRNA"
      /db_xref="taxon:9606"
      /clone="CS0D1039YK15"
      /tissue_type="PLACENTA COT 25-NORMALIZED"
      /note="1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT      184 a   268 c   282 g   194 t      2 others
ORIGIN
Query Match      84.0%; Score 895; DB 13; Length 930;
Best Local Similarity 98.6%; Pred. No. 5.7e-223;
Matches 912; Conservative 0; Mismatches 12; Indels 1; Gaps 1;
QY 1 TCTGAGGGAACAGTACTGCTTCTGGGAATGGTCAAGCTACCGTGGCAGCACAGC 60
DB 927 TTTGAGGGAACAGTACTGCTTCTGGGAATGGTCAAGCTACCGTGGCAGCACAGC 869
QY 61 CTCACCGAGTCGGGTGCTCTCTGCTCCCTCCGCTGGAAATCCATGATCCTGATAGGCAAGTT 120
DB 868 CTCACCGAGTCGGGTGCTCTCTGCTCCCTCCGCTGGAAATCCATGATCCTGATAGGCAAGTT 809
QY 121 TACACGACAGAACCCAGTCCCGAGGACCTGGGCTGGGCAACATTAATTCGCGCG 180
DB 808 TACACGACAGAACCCAGTCCCGAGGACCTGGGCTGGGCAACATTAATTCGCGCG 749
QY 181 AATCTGATGGGATGCCAAGCCCTGCTGACAGCTGCTGAGAACCGCAGGCTGACGTGG 240
DB 748 AATCTGATGGGATGCCAAGCCCTGCTGACAGCTGCTGAGAACCGCAGGCTGACGTGG 689
QY 241 GAGTACTGTGATGTCCTCTCTGCTCCACCTGCGGCTTGAGACAGTACAGCAGGCTCAG 300
DB 688 GAGTACTGTGATGTCCTCTCTGCTCCACCTGCGGCTTGAGACAGTACAGCAGGCTCAG 629
QY 301 TTTCGATCAAGAGGAGGCTTTTCGCGACATCGCTCCACCTCCGAGGCTGCGGCTGCGATC 360
DB 628 TTTCGATCAAGAGGAGGCTTTTCGCGACATCGCTCCACCTCCGAGGCTGCGGCTGCGATC 569
QY 361 TTTCGCAAGCAGCAGGAGTCCGCGGAGAGGCTTCTGTCGGGGGATATCTATCATCAGC 420
DB 568 TTTCGCAAGCAGCAGGAGTCCGCGGAGAGGCTTCTGTCGGGGGATATCTATCATCAGC 509
QY 421 TCCTGCTGGATTCCTCTGCGCGCCCACTGCTTTCCAGGAGAGGTTTCCGCGCCCAACACCTG 480
DB 508 TCCTGCTGGATTCCTCTGCGCGCCCACTGCTTTCCAGGAGAGGTTTCCGCGCCCAACACCTG 449
QY 481 ACGGTGATCTTGGGAGAACATACCGGTGCTTCTGCGGAGGAGGAGAGCAAAATTTGAA 540
DB 448 ACGGTGATCTTGGGAGAACATACCGGTGCTTCTGCGGAGGAGGAGAGCAAAATTTGAA 389
QY 541 GTCGAAATATCATTTGCTCAATGAATTCATGATGATGACCTTACGCAATGACATTCGC 600
DB 388 GTCGAAATATCATTTGCTCAATGAATTCATGATGATGACCTTACGCAATGACATTCGC 329
QY 601 CTGCTGACGCTGAATTCGATTCGCTCCGCTGTGCTCCAGGAGAGCAGGCTGCTCGCACT 660
DB 328 CTGCTGACGCTGAATTCGATTCGCTCCGCTGTGCTCCAGGAGAGCAGGCTGCTCGCACT 269
QY 661 GTGTGCTTCCCGCGGAGACTGTCAGCTGCGGAGACTGGAAGGATGTGAGCTTCTCGGC 720
DB 268 GTGTGCTTCCCGCGGAGACTGTCAGCTGCGGAGACTGGAAGGATGTGAGCTTCTCGGC 209
QY 721 TACGCAAGCATGAGGCTTGTCTCTTCTTATTCGGAGCGGCTGAAGGAGGCTCATGTC 780
DB 208 TACGCAAGCATGAGGCTTGTCTCTTCTTATTCGGAGCGGCTGAAGGAGGCTCATGTC 149
QY 781 AGACTGTACCCATCCAGCGCTGCAATCACAACATTTTACTTAAACAGAACAGTCAACCGAC 840
148 AGACTGTACCCATCCAGCGCTGCAATCACAATTTTACTTAAACAGAACAGTCAACCGAC 89
841 AACATGCTGTGCTGAGACACTCGGAGCGCGCGGCCCGCCCAAGAACTTTCAGCAGCGC 900
88 AACATGCTGTGCTGAGACACTCGGAGCGCGCGGCCCGCCCAAGAACTTTCAGCAGCGC 29
901 TGCCAGGCGGATTCGGAGGCGCC 925
28 TGCCAGGCGGATTCGGAGGCGCC 4
RESULT 2
LOCUS      BUI49958      912 bp      mRNA      linear      EST 03-SEP-2002
DEFINITION AGENCOURT 8137213 Lupski_dorsal_root_ganglion Homo sapiens cDNA
ACCESSION   BUI49958
VERSION     BUI49958.1 GI:22663490
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
REFERENCE   1 (bases 1 to 912)
AUTHORS    NIH-MGC http://mgc.nci.nih.gov/.
TITLE      National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL     Unpublished
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgabs-r@mail.nih.gov
            Tissue Procurement: Dr. James R. Lupski
            CDNA Library Preparation: Life Technologies, Inc.
            CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM13572 row: m column: 16
            High quality sequence start: 18
            High quality sequence stop: 644.
FEATURES
  Location/Qualifiers
    1..912
      /organism="Homo sapiens"
      /mol_type="mRNA"
      /db_xref="taxon:9606"
      /clone="IMAGE:6184119"
      /sex="male"
      /tissue_type="dorsal root ganglia"
      /dev_stage="adult, 36 yr"
      /lab_host="DH10B"
      /clone_lib="Lupski_dorsal_root_ganglion"
      /note="vector: pCMV-SPORT6 (Life Technologies); Site_1:
            NotI; Site 2: SalI; cDNA made by oligo-dr priming.
            Directionally cloned using the following adaptors:
            5'-TGACACCCAGCGTCCG-3' and
            5'-TGACTAGTCTAGATCGGAGCGCGCCCT(15)-3'. Size selected >
            1 kb for average insert length 1.7 kb. This is a primary
            library, non-amplified. Library constructed by Life
            Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
            College of Medicine) and is available through Life
            Technologies."
BASE COUNT      189 a   272 c   271 g   179 t      1 others
ORIGIN
Query Match      80.1%; Score 853; DB 13; Length 912;
Best Local Similarity 98.9%; Pred. No. 5.4e-212;
Matches 901; Conservative 0; Mismatches 5; Indels 5; Gaps 4;
QY 131 AGAACCCAGTGGCCAGGCACTGGG--CCTGGGCAACATTAATTCAGCGGATCCCTGA 188
DB 1 AGAACCCAGTGGCCAGGCACTGGGCGCTGGGCGCAACATTAATTCAGCGGATCCCTGA 60
189 TGGGGATGCCAAGCCCTGTGTGCCACGCTGCTGAAGAACCGCAGGCTGACGTGGGAGTACTG 248
```





```

5', mRNA sequence.
ACCESSION BQ2781177
VERSION BQ2781177.1 GI:20488385
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 993)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCW2043 row: 0 column: 10
High quality sequence stop: 681.
FEATURES
    source
        1..993
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="IMAGE:5805153"
            /tissue_type="teratocarcinoma, cell line"
            /lab_host="DH10B (phage-resistant)"
            /clone_lib="NIH_MGC_109"
            /note="Organ: ovary; Vector: pOTB7; Site 1: EcoRI; Site 2:
            XhoI; cDNA made by oligo-dt priming. Directionally cloned
            into EcoRI/XhoI sites using the following 5' adaptor:
            GGCAGGAG(G). Library constructed by Ling Hong in the
            Laboratory of Gerald M. Rubin (University of California,
            Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
            Superscript II RT (Life Technologies). Note: this is a
            NIH MGC Library."
BASE COUNT 235 a 295 c 269 g 194 t
ORIGIN
Query Match 76.6%; Score 815.8; DB 13; Length 993;
Best Local Similarity 99.6%; Pred. No. 2.9e-202;
Matches 828; Conservative 0; Mismatches 2; Indels 1; Gaps 1;
QY 234 GACGTGGGAGTACTGTGATGTCCTCTCTGCTCCACCTGGCGCTGACAGTACAGCCA 293
DB 1 GACGTGGGAGTACTGTGATGTGCTCTCTCTCTCT-CACCTGGCGCTGACAGTACAGCCA 59
QY 294 GCCTCAGTTTCGATCAAGAGGAGGCTCTTCGCGACATCGCTCCACCCCTGGCAGGC 353
DB 60 GCCTCAGTTTCGATCAAGAGGAGGCTCTTCGCGACATCGCTCCACCCCTGGCAGGC 119
QY 354 TGCCATCTTTCGCAAGCAGAGAGTTCGCCGAGAGCGTTTCTGTGCGGGGGCATACT 413
DB 120 TGCCATCTTTCGCAAGCAGAGAGTTCGCCGAGAGCGTTTCTGTGCGGGGGCATACT 179
QY 414 CATGAGTCTCTGATTTCTCTGCGCGCCACACTGCTTCAGGAGAGGTTTCGCGCCCA 473
DB 180 CATGAGTCTCTGATTTCTCTGCGCGCCACACTGCTTCAGGAGAGGTTTCGCGCCCA 239
QY 474 CCACCTGACGGTGATCTTGGGAGAGGAGTTCGCGGTGCTCTGCGGAGGAGGAGCAGAA 533
DB 240 CCACCTGACGGTGATCTTGGGAGAGGAGTTCGCGGTGCTCTGCGGAGGAGGAGCAGAA 299
QY 534 ATTTGAAGTCGAAAATACATTGTCCATAAGGAATTCGATGATGACATTTACGACAATGA 593
DB 300 ATTTGAAGTCGAAAATACATTGTCCATAAGGAATTCGATGATGACATTTACGACAATGA 359
QY 594 CATTGCGCTGCTGAGCTGAATCGGATTCGTCGCGCTGTGCCAGGAGAGCGGTGGT 653
DB 360 CATTGCGCTGCTGAGCTGAATCGGATTCGTCGCGCTGTGCCAGGAGAGCGGTGGT 419
QY 654 CCGCACTGTGTGCTTCCCGCGGAGACTGCACTGCGGACTGCGGAGTGTGAGCT 713
DB 420 CCGCACTGTGTGCTTCCCGCGGAGACTGCACTGCGGACTGCGGAGTGTGAGCT 479
QY 714 CTCGGCTTACCGCAAGCATGAGGCTTGTCTCTTTTCTATTTCGAGCGGCTGAAGAGGC 773
DB 480 CTCGGCTTACCGCAAGCATGAGGCTTGTCTCTTTTCTATTTCGAGCGGCTGAAGAGGC 539
QY 774 TCATGTGAGCTGTACCCATCCAGCCCTGCACATCAACATTTACTTACAGAACAGT 833
DB 540 TCATGTGAGCTGTACCCATCCAGCCCTGCACATCAACATTTACTTACAGAACAGT 599
QY 834 CACCGACAACATGTGTGTCTGAGACATCTCGGAGCGCGGCGCCAGGCAAACTTCA 893
DB 600 CACCGACAACATGTGTGTCTGAGACATCTCGGAGCGCGGCGCCAGGCAAACTTCA 659
QY 894 CGAGCCTTCCAGGCGGATTCGGAGGCGCCCTGCTGTCTGAAAGTATGCGCGCATGAC 953
DB 660 CGAGCCTTCCAGGCGGATTCGGAGGCGCCCTGCTGTCTGAAAGTATGCGCGCATGAC 719
QY 954 TTTGCTGGGCATCATCAGTGGGCGCTGGGCTGTGGACAGAGGATGTCGCGGTGTGTA 1013
DB 720 TTTGCTGGGCATCATCAGTGGGCGCTGGGCTGTGGACAGAGGATGTCGCGGTGTGTA 779
QY 1014 CACAAAGGTTACCACTACCTAGACTGATTCGTGACAACTGCGACCGTG 1064
DB 780 CACCAAGGTTACCACTACCTAGACTGATTCGTGACAACTGCGACCGTG 830
RESULT 6
BO690978 883 bp mRNA linear EST 15-JUL-2002
LOCUS AGENCOURT_8046707 NIH_MGC_110 Homo sapiens cDNA clone IMAGE:6209120
DEFINITION 5', mRNA sequence.
ACCESSION BQ690978
VERSION BQ690978.1 GI:21816294
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 883)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM2367 row: 0 column: 09
High quality sequence stop: 658.
FEATURES
    source
        1..883
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="IMAGE:6209120"
            /tissue_type="ductal carcinoma, cell line"
            /lab_host="DH10B (phage-resistant)"
            /clone_lib="NIH_MGC_110"
            /note="Organ: pancreas; Vector: pOTB7; Site 1: XhoI;
            Site 2: EcoRI; cDNA made by oligo-dt priming.
            Directionally cloned into EcoRI/XhoI sites using the
            following 5' adaptor: GGCAGGAG(G). Library constructed by
            Ling Hong in the laboratory of Gerald M. Rubin (University
            of California, Berkeley) using ZAP-cDNA synthesis kit

```

(Stratagene) and Superscript II RT (Life Technologies).  
Note: this is a NIH\_MGC Library."

BASE COUNT 184 a 265 c 259 g 175 t

ORIGIN

Query Match 76.2%; Score 811.2; DB 13; Length 883;  
Best Local Similarity 99.0%; Pred. No. 4.5e-201;  
Matches 816; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

```
QY 112 GCACAGGTTTACACAGCAGAACCCAGTGCAGGCACTGGGCTGGGCAAACTAAT 171
Db 1 GCACAGGTTTACACAGCAGAACCCAGTGCAGGCACTGGGCTGGGCAAACTAAT 60

QY 172 TACTGCCGAATCCTGATGGGATGCCAAGCCCTGGTGCACGCTGCTGAAGAACCGCAGG 231
Db 61 TACTGCCGAATCCTGATGGGATGCCAAGCCCTGGTGCACGCTGCTGAAGAACCGCAGG 120

QY 232 CTGACGTGGAGTACTGTGATGTGCCCTCTCTCTCCACTGCTGGGCTGAGACAGTACAGC 291
Db 121 CTGACGTGGAGTACTGTGATGTGCCCTCTCTCTCCACTGCTGGGCTGAGACAGTACAGC 180

QY 292 CAGCCTCAGTTTCGATCAAGAGGGGCTCTTCGCGGACATCGCTCCACCCCTGGCAG 351
Db 181 CAGCCTCAGTTTCGATCAAGAGGGGCTCTTCGCGGACATCGCTCCACCCCTGGCAG 240

QY 352 GCTGCCATCTTTGCCAAGCACAGGAGTGCAGGAGCGGTTCTGTGGGGGGCATA 411
Db 241 GCTGCCATCTTTGCCAAGCACAGGAGTGCAGGAGCGGTTCTGTGGGGGGCATA 300

QY 412 CTCATCAGTCTCTGTGATCTCTCTGCGGCCACTGCTCCAGAGAGGTTTCGCGCC 471
Db 301 CTCATCAGTCTCTGTGATCTCTCTGCGGCCACTGCTCCAGAGAGGTTTCGCGCC 360

QY 472 CACCACTCAGGTCATCTTGGGCAAGACATACCGGTCCTCCCTGGCGAGGAGGAGC 531
Db 361 CACCACTCAGGTCATCTTGGGCAAGACATACCGGTCCTCCCTGGCGAGGAGGAGC 420

QY 532 AAATTTGAAGTCGAAATATACATTTGCATAGGAATTCGATGATGACATTTACGACAT 591
Db 421 AAATTTGAAGTCGAAATATACATTTGCATAGGAATTCGATGATGACATTTACGACAT 480

QY 592 GACATGCGCTGCTCAGCTGAAATCGGATTCGTCCTCCGCTGTGCCAGAGAGCAGCTG 651
Db 481 GACATGCGCTGCTCAGCTGAAATCGGATTCGTCCTCCGCTGTGCCAGAGAGCAGCTG 540

QY 652 GTCCGCACTGTGTGCTTCCCGCGGACCTGACGCTGCCGACTGACGAGTGTGAG 711
Db 541 GTCCGCACTGTGTGCTTCCCGCGGACCTGACGCTGCCGACTGACGAGTGTGAG 600

QY 712 CTCTCCGCTACGGCAAGCATGAGGCTTGTCTCTCTTTCTATTTCGGAGCGCTGAAGGAG 771
Db 601 CTCTCCGCTACGGCAAGCATGAGGCTTGTCTCTCTTTCTATTTCGGAGCGCTGAAGGAG 660

QY 772 GCTCATGTGAGCTGTACCCATCCAGCGCTGCACATCACAACTTTACTTAAACAGACA 831
Db 661 GCTCATGTGAGCTGTACCCATCCAGCGCTGCACATCACAACTTTACTTAAACAGACA 720

QY 832 GTACCGCAACATGCTGTGTGCTGGACACTCGAGCGGGGCGGCCCGCAAACTTG 891
Db 721 GTACCGCAACATGCTGTGTGCTGGACACTCGAGCGGGGCGGCCCGCAAACTTG 780

QY 892 CACGACGCTGCGAGGGGATTCGGAGGCGCCCTCTGTGTGTCT 935
Db 781 CACGACGCTGCGAGGGGATTCGGAGGCGCCCTCTGTGTGTCT 824
```

RESULT 7  
BX389608  
LOCUS  
DEFINITION  
BX389608 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA  
clone CS0D1039YK15 5-PRIME, mRNA sequence.  
ACCESSION  
BX389608  
VERSION  
BX389608.1 GI:30463469

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE  
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

TITLE 1 (bases 1 to 918)

JOURNAL Li, W.B., Gruber, C., Jessee, J. and Polayes, D.

COMMENT Full-length cDNA libraries and normalization

Unpublished

Contact: Genoscope

Genoscope - Centre National de Sequencage

BP 191 91006 Evry cedex - France

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr

was normalized. Library was constructed by Life Technologies, a

division of Invitrogen. This sequence belongs to sequence cluster

4498.r For more information about this cluster, see

http://www.genoscope.cns.fr/

cgi-bin/cluster.cgi?seq=CS0BAF025ZG09\_AF02376\_2&cluster=4498.r.

Contact : Feng Liang Email : fliang@lifetech.com URL :

http://fulllength.invitrogen.com/Invitrogen Corporation 1600

Paradise Avenue Genoscope sequence ID : CS0BAF025ZG09\_AF02376\_2.

FEATURES

source

1. 918

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="CS0D1039YK15"

/tissue\_type="PLACENTA COT 25-NORMALIZED"

/clone\_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"

/note="1st strand cDNA was primed with a NotI-oligo (dT)

primer. Five prime end enriched, double-strand cDNA was

digested with Not I and cloned into the Not I and EcoR V

sites of the pCMVSPORT 6 vector. Library was normalized."

BASE COUNT 189 a 273 c 264 g 186 t 6 others

ORIGIN

Query Match

Best Local Similarity 75.8%; Score 807.6; DB 13; Length 918;

Matches 855; Conservative 0; Mismatches 29; Indels 3; Gaps 3;

QY 170 ATTACTCCGGAATCTGTATGGGATCCAAAGCCCTGGTCCACGCTGTGAAGAACCGCA 229

Db 17 ATTACTCCGGAATCTGTATGGGATCCAAAGCCCTGGTCCACGCTGTGAAGAACCGCA 76

QY 230 GGCTGACGTGGGAGTACTGTGATGTGCTCTCTCCCTCCACTGCGGCTGAGACAGTACA 289

Db 77 GGCTGACGTGGGAGTACTGTGATGTGCTCTCTCCCTCCACTGCGGCTGAGACAGTACA 136

QY 290 GCCAGCCTCAGTTTCGATCAAGAGGAGGCTCTTCGCGGACATCGCTCCACCCCTGGC 349

Db 137 GCCAGCCTCAGTTTCGATCAAGAGGAGGCTCTTCGCGGACATCGCTCCACCCCTGGC 196

QY 350 AGGCTGCCATCTTTGCCAAGCACAGGAGGTGCCTCCGAGAGCGGTTCCTGTGCGGGGCA 409

Db 197 AGGCTGCCATCTTTGCCAAGCACAGGAGGTGCCTCCGAGAGCGGTTCCTGTGCGGGGCA 256

QY 410 TACTCATCAGTCTCTGTGATTTCTCTGCGGCCACTGCTCCAGAGAGGTTTCGCG 469

Db 257 TACTCATCAGTCTCTGTGATTTCTCTGCGGCCACTGCTCCAGAGAGGTTTCGCG 316

QY 470 CCACACCTCAGCGGTGATCTTGGGCAAGACATACCGGGTGGTCCCTGGCGAGGAGGAGC 529

Db 317 CCACACCTCAGCGGTGATCTTGGGCAAGACATACCGGGTGGTCCCTGGCGAGGAGGAGC 376

QY 530 AGAAATTTGAAGTCGAAAAATAATTGTCATAGGAATTCGATGATGACACTTACGACA 589

Db 377 AGAAATTTGAAGTCGAAAAATAATTGTCATAGGAATTCGATGATGACACTTACGACA 436

QY 590 ATGACATTCGCTGCTGCAGCTGAAATCGATTCCTCCCTGTCGCCAGGAGGAGGAGC 649

Db 437 ATGACATTCGCTGCTGCAGCTGAAATCGATTCCTCCCTGTCGCCAGGAGGAGGAGC 496

QY 650 TGGTCCGCACTGTGTGCTTCCCTCCCGCGGACCTGACGCTGCCGAGTGTG 709





Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE  
AUTHORS  
1 (bases 1 to 888)  
NTH-MGC <http://mgc.nci.nih.gov/>.

**TITLE** National Institutes of Health, Mammalian Gene Collection (MGC)  
**JOURNAL** Unpublished

UNPUBLISHED  
COMMENT  
Contact: Robert Strausberg, Ph.D.  
Email: [ccarbs-r@mail.nih.gov](mailto:ccarbs-r@mail.nih.gov)

**Tissue Procurement:** ATCC  
cDNA Library Preparation: Rubin Laboratory  
Email: [cgapops-remail@nih.gov](mailto:cgapops-remail@nih.gov)

CDNA Library Prepared by: The I.M.A.G.E. Consortium (LLNL)  
CDNA Sequencing by: Agencourt Bioscience Corporation

DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be found through the T.M.C. Corporation/Invitrogen website.

found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>

Plate: LLM2316 row: e column: 21  
High quality sequence stop: 630.

FEATURES  
Location/Qualifiers  
1..888  
source

```
/organism="Homo sapiens"  
/mol_type="mRNA"
```

```
/db_xref="taxon:9606"  
/clone="IMAGE:6084476"
```

```
/tissue type="ductal carcinoma, cell line"  
/lab host="DH10B (phage-resistant)"
```

```

/clone_lib="NIH_MGC_110"
/note="Organ: pancreas; Vector: pOTB7; Site 1: XhoI;

```

Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the

following 5' adaptor: GGCACGAG(G). Library constructed using EcoRI/NotI sites using the directionally cloned into EcoRI/NotI sites using the

of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies) working in the laboratory of Gerald M. Rubin (University of California, Berkeley).

BASE COUNT 181 a 265 c 270 c - 172 +

BASE COUNT	181 a	265 c	270 g	172 c
ORIGIN				

Query Match 73.5%; Score 782.8; DB 13; Length 888;

Best Local Similarity 97.2%; Pred. No. 1.2e-193;  
Matches 861; Conservative 0; Mismatches 17; Indels 8; Gaps

QY 112 GGCAAGTTTACACAGCACAGAACCCACGTGCCAGGCCCTGGGCAACATAAT

Db 1 GGCAAGTTTACACAGCAGAACCCCACTGCCAGGCACTGGGCGCTGGGCAACACATAAT

QY 172 TACTGCCGAATCCTCATGGGATGCCAGCCCTGGTGCCACGTGCTGAAGAACCGCAGG

61 TACTGCCGAATCTTGATGGGAGTGCACAGCCCTGTGTGCCACGTGCTGAAGAAACCGCAGG



Gruber (Invitrogen). Research Genetics tracking code  
023. Note: this is a NIH MGC Library.  
204 a 272 c 256 g 184 t

BASE COUNT 204 a 272 c 256 g 184 t

ORIGIN

Query Match 72.1%; Score 767.4; DB 12; Length 916;  
Best Local Similarity 99.7%; Pred. No. 1.3e-189;  
Matches 779; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 286 TACAGCCAGCCTCAGTTTCGCATCAAGAGAGGGCTCTTCGCGAGATCGCTCCACACCC 345  
DB 1 TACAGCCAGCCTCAGTTTCGCATCAAGAGAGGGCTCTTCGCGAGATCGCTCCACACCC 60

QY 346 TGGCAGGCTGCCATCTTTGCCAAGACAGAGAGGTGCCCGAGAGCGGTTCTGTGCGG 405  
DB 61 TGGCAGGCTGCCATCTTTGCCAAGACAGAGAGGTGCCCGAGAGCGGTTCTGTGCGG 120

QY 406 GGCATACATCAGCTCCTGCTGGATCTCTCTGCGGCCACTGCTTCCAGAGAGGTTT 465  
DB 121 GGCATACATCAGCTCCTGCTGGATCTCTCTGCGGCCACTGCTTCCAGAGAGGTTT 180

QY 466 CGCGCCACACCTGACGGTGATCTTGGGAGAGATACCGGGTGGTCCCTGGCGAGGAG 525  
DB 181 CGCGCCACACCTGACGGTGATCTTGGGAGAGATACCGGGTGGTCCCTGGCGAGGAG 240

QY 526 GAGCAGAAATTGAAGTGGAAATATATTTGCCATTAAGGAATTCGATGATGACATTTAC 585  
DB 241 GAGCAGAAATTGAAGTGGAAATATATTTGCCATTAAGGAATTCGATGATGACATTTAC 300

QY 586 GACATGACATTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 645  
DB 301 GACATGACATTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360

QY 646 AGCGTGTTCGACATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 705  
DB 361 AGCGTGTTCGACATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420

QY 706 TGTGAGCTCTCGGCTACGGCAAGCATGAGCGCTTGTCTCTCTTCTATTCGGAGCGGCTG 765  
DB 421 TGTGAGCTCTCGGCTACGGCAAGCATGAGCGCTTGTCTCTCTTCTATTCGGAGCGGCTG 480

QY 766 AAGGAGGCTCATGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTG 825  
DB 481 AAGGAGGCTCATGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTG 540

QY 826 AGAAGAGTCAACGACATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 885  
DB 541 AGAAGAGTCAACGACATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600

QY 886 AACTTGCACGAGCTGCGAGGCGGATTCGGAGGCGGCTGCTGCTGCTGCTGCTGCTG 945  
DB 601 AACTTGCACGAGCTGCGAGGCGGATTCGGAGGCGGCTGCTGCTGCTGCTGCTGCTG 660

QY 946 GCGATGATCTTGGTGGGATCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1005  
DB 661 GCGATGATCTTGGTGGGATCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720

QY 1006 GGTGTGTACAAAGGTTTACCACTACCTAGAC-TGGATTGCTGACACATGCGACCGTG 1064  
DB 721 GGTGTGTACAAAGGTTTACCACTACCTAGAC-TGGATTGCTGACACATGCGACCGTG 780

QY 1065 A 1065  
DB 781 A 781

RESULT 11  
BQ927899  
LOCUS  
DEFINITION AGNCOURT 8775988 NIH\_MGC\_18 Homo sapiens cDNA clone IMAGE:6373489  
5', mRNA sequence.  
ACCESSION BQ927899  
VERSION BQ927899.1 GI:22342930

KEYWORDS EST. Homo sapiens (human)  
SOURCE Homo sapiens  
ORGANISM Homo sapiens  
REFERENCE 1 (bases 1 to 950)  
AUTHORS NIH-MGC http://mhc.nci.nih.gov/  
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)  
COMMENT Unpublished  
Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov  
Tissue Procurement: DCTD/DTP/Gazdar  
cDNA Library Preparation: Rubin Laboratory  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LICM2550 row: p column: 02  
High quality sequence start: 35  
High quality sequence stop: 684.

FEATURES  
source  
1..950  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:6373489"  
/tissue\_type="large cell carcinoma"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH\_MGC\_18"  
/note="Organ: lung; Vector: pOTB7; Site: 1: XhoI; Site 2:  
EcoRI; cDNA made by oligo-dr priming. Directionally cloned  
into EcoRI/XhoI sites using the following 5' adaptor:  
GGCAGCAG(G). Library constructed by Ling Hong in the  
laboratory of Gerald M. Rubin (University of California,  
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and  
Superscript II RT (Life Technologies). Note: this is a  
NIH MGC Library."  
BASE COUNT 207 a 283 c 279 g 181 t  
ORIGIN

Query Match 71.8%; Score 765.2; DB 13; Length 950;  
Best Local Similarity 98.1%; Pred. No. 4.9e-189;  
Matches 806; Conservative 0; Mismatches 13; Indels 3; Gaps 3;

QY 210 CCAGCTGCTGAAGAACCGCAGCGCTGAGTGGAGTACTGTGATGTCCTCTCTCCAC 269  
DB 43 CCCCCTGCTGAAGAACCGCAGCGCTGAGTGGAGTACTGTGATGTCCTCTCTCCAC 102

QY 270 CTGCGGCTGAGACAGTACAGCAGCCTCAGTTTCGCATCAAGAGAGGCTCTTCGCCGA 329  
DB 103 CTGCGGCTGAGACAGTACAGCAGCCTCAGTTTCGCATCAAGAGAGGCTCTTCGCCGA 162

QY 330 CATCGCTCCACCCCTGGCAGGCTGCCATCTTTGCCAAGCAGAGAGGTGCCCGGAGA 389  
DB 163 CATCGCTCCACCCCTGGCAGGCTGCCATCTTTGCCAAGCAGAGAGGTGCCCGGAGA 222

QY 390 GCGGTTCTGTGCGGGGCTACTCATCAGCTCTGCTGGATTCTCTCTGCGGCCACTG 449  
DB 223 GCGGTTCTGTGCGGGGCTACTCATCAGCTCTGCTGGATTCTCTCTGCGGCCACTG 282

QY 450 CTTCCAGAGAGG-TTTCCGCCCCCACCCTGAGCGGTGATCTTGGGAGAGATACACCGG 508  
DB 283 CTTCCAGAGAGGTTTCTTCCGCCCCCACCCTGAGCGGTGATCTTGGGAGAGATACACCGG 342

QY 509 TGGTCCCTGGCAGGAGGAGCAGAAATTGAAGTCGAAAAATACATTGTTCATAGGAAT 568  
DB 343 TGGTCCCTGGCAGGAGGAGCAGAAATTGAAGTCGAAAAATACATTGTTCATAGGAAT 402

QY 569 TCGATGATGACACTTACGACATGACATTCGCTGCTGCGAGCTGAATTCGATTCGCTCC 628  
DB 403 TCGATGATGACACTTACGACATGACATTCGCTGCTGCGAGCTGAAATCGGATTCGCTCC 462





XhoI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCAGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH MGC Library."

BASE COUNT 215 a 293 c 279 g 187 t  
ORIGIN  
Query Match 65.2%; Score 694.4; DB 13; Length 974;  
Best Local Similarity 97.8%; Pred. No. 1.6e-170;  
Matches 757; Conservative 0; Mismatches 11; Indels 6; Gaps 5;  
QY 276 CTGGAGAGTACAGCAGCCTCAGTTTCGATCAAGAGGGCTCTTCGGCGCATCGC 335  
DB 19 CTGGAGAGTACAGCAGCCTCAGTTTCGATCAAGAGGGCTCTTCGGCGCATCGC 78  
QY 336 CTCCACCCCTGGCAGGCTGCCATCTTTGCCAAGCAGAGGAGTCCGCCGGAGAGCGGTT 395  
DB 79 CTCCACCCCTGGCAGGCTGCCATCTTTGCCAAGCAGAGGAGTCCGCCGGAGAGCGGTT 138  
QY 396 CTGTGCGGGGCATACATCAGCTCTGTGGATTCTCTGCGGCCCACTGCTTCCA 455  
DB 139 CTGTGCGGGGCATACATCAGCTCTGTGGATTCTCTGCGGCTC-CTGCTTCCA 197  
QY 456 GGAGAGGTTTCGCCCCACCACTGACGGTATCTTGGCGAGACATACCGGGTGTCCC 515  
DB 198 GGAGAGGTTTCGCCCCACCACTGACGGTATCTTGGCGAGACATACCGGGTGTCCC 257  
QY 516 TGGCGAGGAGGACGAAATTTGAAGTCGAAATATACATTTGCCATAGGAATTCGATGA 575  
DB 258 TGGCGAGGAGGACGAAATTTGAAGTCGAAATATACATTTGCCATAGGAATTCGATGA 317  
QY 576 TGACACTTACGACATATGATTCGCTCTGCGCTGAAATTCGATTCGCTCCCGCTGTC 635  
DB 318 TGACACTTACGACATATGATTCGCTCTGCGCTGAAATTCGATTCGCTCCCGCTGTC 377  
QY 636 CCAGAGAGCAGCGTGTGCGCACTGTGCTTCCCGGGGAGCCTGCGAGCTGCCGGA 695  
DB 378 CCAGAGAGCAGCGTGTGCGCACTGTGCTTCCCGGGGAGCCTGCGAGCTGCCGGA 437  
QY 696 CTGGACGAGTGTGAGCTCTCGGCTACGGAAGCATGAGGCTGTCTCTTCTTATTC 755  
DB 438 CTGGACGAGTGTGAGCTCTCGGCTACGGAAGCATGAGGCTGTCTCTTCTTATTC 497  
QY 756 GGAGCGGTGAAGGAGGCTCATGTGAGTGTACCCATCCAGCGCTGCACATCAACA 815  
DB 498 GGAGCGGTGAAGGAGGCTCATGTGAGTGTACCCATCCAGCGCTGCACATCAACA 557  
QY 816 TTTACTTTAAGAACAGTACCGACAATGCTGTGTGTCGAGACACTCGAGGGGGG 875  
DB 558 TTTACTTTAAGAACAGTACCGACAATGCTGTGTGTCGAGACACTCGAGGGGGG 617  
QY 876 GCCCAGGCAAACTTGCACGAGCCTGCGAGGGGATTCGGAGG-CCCCCTGGTGTCT 934  
DB 618 GCCCAGGCAAACTTGCACGAGCCTGCGAGGGGATTCGGAGGCCCCCTGGTGTCT 677  
QY 935 TGAACGATGGCGCAT-GACTTTGGTGGGCATCATCAGCT-GGGGCTTGGCTGTGGACA 992  
DB 678 TGAACGATGGCGCATGAGCTTTGGTGGGCATCATCAGCTGGGGGCTGGGCTGTGGACA 737  
QY 993 GAAGGATGTC-GGGTGTGTACAAAGGTTTACCACTACCTAGACTGGAAT 1044  
DB 738 GAAGGATGTCGGGGGGGTACCCCAAGGGTTTACCACTACCTAAACTGGAAT 791

RESULT 15  
BE732704  
LOCUS 601571366F1 NIH\_MGC\_21 Homo sapiens cDNA clone IMAGE:392560 5',  
DEFINITION mRNA sequence.  
ACCESSION BE732704

VERSION  
KEYWORDS  
SOURCE

BE732704.1 GI:10146696  
EST.  
Homo sapiens (human)

ORGANISM

Homo sapiens

REFERENCE

1 (bases 1 to 709)  
NIH-MGC http://mgi.nci.nih.gov/  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished

AUTHORS

TITLE

JOURNAL

COMMENT

Contact: Robert Strausberg, Ph.D.  
Email: cgabs-r@mail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: Ling Hong/Rubin Laboratory

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov

Plate: LLCM752 row: d column: 17

High quality sequence start: 3

High quality sequence stop: 701.

Location/Qualifiers

FEATURES

source

1. 709

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clones="IMAGE:3925600"

/tissue\_type="choriocarcinoma"

/lab\_host="DH10B (phage-resistant)"

/clone\_lib="NIH MGC 21"

/note="Organ: placenta; Vector: pOTB7; Site 1: XhoI;

Site 2: EcoRI; cDNA made by oligo-dT priming.

Directionally cloned into EcoRI/XhoI sites using the

following 5' adaptor: GGCAGAG(G). Size-selected >500bp

for average insert size 1.8kb. Library constructed by

Ling Hong in the laboratory of Gerald M. Rubin (University

of California, Berkeley) using ZAP-cDNA synthesis kit

(Stratagene) and Superscript II RT (Life Technologies)."

BASE COUNT 147 a 219 c 203 g 140 t

ORIGIN

Query Match 65.1%; Score 693.4; DB 10; Length 709;

Best Local Similarity 99.7%; Pred. No. 2.5e-170;

Matches 705; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 197 CCAAGCCCTGGTGCACAGTGTGAAGAACCGCAGGCTGACGTGGAGTACTGTGATGTC 256

DB 1 CCAAGCCCTGGTGCACAGTGTGAAGAACCGCAGGCTGACGTGGAGTACTGTGATGTC 60

QY 257 CCTCTCTCCACCTGGCGCTGAGACAGTACAGCCAGCCTCAGTTTCGCATCAAGGAG 316

DB 61 CCTCTCTCCACCTGGCGCTGAGACAGTACAGCCAGCCTCAGTTTCGCATCAAGGAG 120

QY 317 GGCTCTTTCGCGCATCGCTCCACCCCTGGCAGGCTGCCATCTTTGCCAAGCAGGAG 376

DB 121 GGCTCTTTCGCGCATCGCTCCACCCCTGGCAGGCTGCCATCTTTGCCAAGCAGGAG 180

QY 377 GGTCGCCCGGAGAGCGGTTCTGTGCGGGGAGCATCATCAGCTCTCTGCTGGATTCTCT 436

DB 181 GGTCGCCCGGAGAGCGGTTCTGTGCGGGGAGCATCATCAGCTCTCTGCTGGATTCTCT 240

QY 437 CTGCGCCGCCACTGCTCCAGGAGAGGTTTCGCCCCACCCAGCCTGACGCTGATCTTGGGCA 496

DB 241 CTGCGCCGCCACTGCTCCAGGAGAGGTTTCGCCCCACCCAGCCTGACGCTGATCTTGGGCA 300

QY 497 GAAATACCCGGTGTCTCCCTGGCGAGGAGGAGCAGAAATTTGAAGTCGAAATATCATTTG 556

DB 301 GAAATACCCGGTGTCTCCCTGGCGAGGAGGAGCAGAAATTTGAAGTCGAAATATCATTTG 360

QY 557 TCCATAAGGAATTCGATGATGACATTTACGACAATGATTCGCTGCTGCTGAGTGAAT 616

DB 361 TCCATAAGGAATTCGATGATGACATTTACGACAATGATTCGCTGCTGCTGAGTGAAT 420

Search completed: August 19, 2003, 08:29:17  
Job time : 3116.64 secs

THIS PAGE BLANK (USP 10)

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model  
Run on: August 18, 2003, 19:53:48 ; Search time 75.5794 Seconds  
(without alignments)  
9743.047 Million cell updates/sec  
Title: US-09-987-455-1  
Perfect score: 18  
Sequence: 1 tctgagggaacagtgc 18  
Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0  
Searched: 2888711 seqs, 20454813386 residues  
Total number of hits satisfying chosen parameters: 5777422  
Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl.\*

- 1: gb.ba.\*
- 2: gb.htg.\*
- 3: gb.in.\*
- 4: gb.om.\*
- 5: gb.ov.\*
- 6: gb.pat.\*
- 7: gb.ph.\*
- 8: gb.pl.\*
- 9: gb.pr.\*
- 10: gb.ro.\*
- 11: gb.sts.\*
- 12: gb.sy.\*
- 13: gb.un.\*
- 14: gb.vi.\*
- 15: em.ba.\*
- 16: em.fun.\*
- 17: em.hum.\*
- 18: em.in.\*
- 19: em.mu.\*
- 20: em.om.\*
- 21: em.or.\*
- 22: em.ov.\*
- 23: em.pat.\*
- 24: em.ph.\*
- 25: em.pl.\*
- 26: em.ro.\*
- 27: em.sts.\*
- 28: em.un.\*
- 29: em.vi.\*
- 30: em.htg.hum.\*
- 31: em.htg.lnv.\*
- 32: em.htg.other.\*
- 33: em.htg.mus.\*
- 34: em.htg.pln.\*
- 35: em.htg.rtd.\*
- 36: em.htg.mam.\*
- 37: em.htg.vrt.\*
- 38: em.sy.\*
- 39: em.htgo.hum.\*
- 40: em.htgo.mus.\*
- 41: em.htgo.other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	18	100.0	18	6	AX430833	AX430833 Sequence
2	18	100.0	18	6	AX431267	AX431267 Sequence
3	18	100.0	42	6	AX430834	AX430834 Sequence
4	18	100.0	42	6	AX431288	AX431288 Sequence
5	18	100.0	952	6	IO9039	IO9039 Sequence 2
6	18	100.0	1065	6	A27447	A27447 DNA sequence
7	18	100.0	1065	6	AR059995	AR059995 Sequence
8	18	100.0	1065	6	AX431270	AX431270 Sequence
9	18	100.0	1065	6	AX431273	AX431273 Sequence
10	18	100.0	1065	6	E01943	E01943 Synthetic D
11	18	100.0	1068	6	A27431	A27431 DNA sequenc
12	18	100.0	1068	6	A27433	A27433 DNA sequenc
13	18	100.0	1068	6	A27439	A27439 DNA sequenc
14	18	100.0	1068	6	A27445	A27445 DNA sequenc
15	18	100.0	1068	6	AR059987	AR059987 Sequence
16	18	100.0	1068	6	AR059988	AR059988 Sequence
17	18	100.0	1068	6	AR059991	AR059991 Sequence
18	18	100.0	1068	6	AR059994	AR059994 Sequence
19	18	100.0	1068	6	E01935	E01935 Synthetic D
20	18	100.0	1068	6	E01936	E01936 Synthetic D
21	18	100.0	1068	6	E01939	E01939 Synthetic D
22	18	100.0	1068	6	E01942	E01942 Synthetic D
23	18	100.0	1068	6	E05897	E05897 DNA sequenc
24	18	100.0	1068	6	E05898	E05898 DNA sequenc
25	18	100.0	1128	6	AX431268	AX431268 Sequence
26	18	100.0	1128	6	AX431271	AX431271 Sequence
27	18	100.0	1170	6	A27727	A27727 DNA sequenc
28	18	100.0	1170	6	A27729	A27729 DNA sequenc
29	18	100.0	1170	6	AR059997	AR059997 Sequence
30	18	100.0	1170	6	AR059998	AR059998 Sequence
31	18	100.0	1170	6	E01945	E01945 Synthetic D
32	18	100.0	1170	6	E01946	E01946 Synthetic D
33	18	100.0	1185	9	AF260825	AF260825 Homo sapi
34	18	100.0	1314	6	A27435	A27435 DNA sequenc
35	18	100.0	1314	6	A27437	A27437 DNA sequenc
36	18	100.0	1314	6	A27441	A27441 DNA sequenc
37	18	100.0	1314	6	AR059989	AR059989 Sequence
38	18	100.0	1314	6	AR059990	AR059990 Sequence
39	18	100.0	1314	6	AR059992	AR059992 Sequence
40	18	100.0	1314	6	AR059993	AR059993 Sequence
41	18	100.0	1314	6	E01937	E01937 Synthetic D
42	18	100.0	1314	6	E01938	E01938 Synthetic D
43	18	100.0	1314	6	E01940	E01940 Synthetic D
44	18	100.0	1314	6	E01941	E01941 Synthetic D
45	18	100.0	1341	6	A30593	A30593 DNA for tis

ALIGNMENTS

RESULT 1  
AX430833  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL

AX430833  
Sequence 5 from Patent WO0240696.  
AX430833  
AX430833.1 GI:21655910  
synthetic construct  
synthetic construct  
artificial sequences.  
1  
Goetz, F., Werner, R.G., Tayapiwatana, C., Manosroi, A. and Manosroi, J.  
Methods for large scale protein production in prokaryotes  
Patent: WO 0240696-A 5 23-MAY-2002;  
BOEHRINGER INGELHEIM INT (DE)

AX430833  
Sequence 5 from Patent WO0240696.  
AX430833  
AX430833.1 GI:21655910  
synthetic construct  
synthetic construct  
artificial sequences.  
1  
Goetz, F., Werner, R.G., Tayapiwatana, C., Manosroi, A. and Manosroi, J.  
Methods for large scale protein production in prokaryotes  
Patent: WO 0240696-A 5 23-MAY-2002;  
BOEHRINGER INGELHEIM INT (DE)

AX430833  
Sequence 5 from Patent WO0240696.  
AX430833  
AX430833.1 GI:21655910  
synthetic construct  
synthetic construct  
artificial sequences.  
1  
Goetz, F., Werner, R.G., Tayapiwatana, C., Manosroi, A. and Manosroi, J.  
Methods for large scale protein production in prokaryotes  
Patent: WO 0240696-A 5 23-MAY-2002;  
BOEHRINGER INGELHEIM INT (DE)

```

FEATURES
  source
    Location/Qualifiers
      1..18
      /organism="synthetic construct"
      /mol_type="genomic DNA"
      /db_xref="taxon:32630"
      /note="coding sequence of the N-terminal part of K2S
      molecule"
    6 a 3 c 6 g 3 t

Query Match
  Best Local Similarity 100.0%; Score 18; DB 6; Length 18;
  Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCTGAGGGAACAGTGAC 18
    |||||
Db 1 TCTGAGGGAACAGTGAC 18

RESULT 2
LOCUS AX431267 18 bp DNA linear PAT 28-JUN-2002
DEFINITION Sequence 1 from Patent WO0240650.
ACCESSION AX431267
VERSION AX431267.1 GI:21656149
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
          artificial sequences.
REFERENCE
  1 Goetz,F., Werner,R.G., Tayapiwatana,C., Manosroi,A. and Manosroi,J.
  TITLE Methods for large scale production of recombinant dna-derived tpa
  or k2s molecules
  JOURNAL Patent: WO 0240650-A 1 23-MAY-2002;
  BOEHRINGER INGELHEIM INT (DE)
FEATURES
  source
    Location/Qualifiers
      1..18
      /organism="synthetic construct"
      /mol_type="genomic DNA"
      /db_xref="taxon:32630"
      /note="coding sequence of N-terminal part of K2S protein"
    6 a 3 c 6 g 3 t

Query Match
  Best Local Similarity 100.0%; Score 18; DB 6; Length 18;
  Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCTGAGGGAACAGTGAC 18
    |||||
Db 1 TCTGAGGGAACAGTGAC 18

RESULT 3
LOCUS AX430834 42 bp DNA linear PAT 28-JUN-2002
DEFINITION Sequence 6 from Patent WO0240696.
ACCESSION AX430834
VERSION AX430834.1 GI:21655911
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
          artificial sequences.
REFERENCE
  1 Goetz,F., Werner,R.G., Tayapiwatana,C., Manosroi,A. and Manosroi,J.
  TITLE Methods for large scale protein production in prokaryotes
  JOURNAL Patent: WO 0240696-A 6 23-MAY-2002;
  BOEHRINGER INGELHEIM INT (DE)
FEATURES
  source
    Location/Qualifiers
      1..42
      /organism="synthetic construct"
      /mol_type="genomic DNA"
      /db_xref="taxon:32630"

```

```

BASE COUNT 10 a 9 c 19 g 4 t
ORIGIN
Query Match
  Best Local Similarity 100.0%; Score 18; DB 6; Length 42;
  Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCTGAGGGAACAGTGAC 18
    |||||
Db 25 TCTGAGGGAACAGTGAC 42

RESULT 4
LOCUS AX431288 42 bp DNA linear PAT 28-JUN-2002
DEFINITION Sequence 22 from Patent WO0240650.
ACCESSION AX431288
VERSION AX431288.1 GI:21656157
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
          artificial sequences.
REFERENCE
  1 Goetz,F., Werner,R.G., Tayapiwatana,C., Manosroi,A. and Manosroi,J.
  TITLE Methods for large scale production of recombinant dna-derived tpa
  or k2s molecules
  JOURNAL Patent: WO 0240650-A 22 23-MAY-2002;
  BOEHRINGER INGELHEIM INT (DE)
FEATURES
  source
    Location/Qualifiers
      1..42
      /organism="synthetic construct"
      /mol_type="genomic DNA"
      /db_xref="taxon:32630"
      /note="primer"
    10 a 9 c 19 g 4 t

BASE COUNT 10 a 9 c 19 g 4 t
ORIGIN
Query Match
  Best Local Similarity 100.0%; Score 18; DB 6; Length 42;
  Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCTGAGGGAACAGTGAC 18
    |||||
Db 25 TCTGAGGGAACAGTGAC 42

RESULT 5
LOCUS I09039 952 bp DNA linear PAT 02-DEC-1994
DEFINITION Sequence 2 from Patent WO 8808878.
ACCESSION I09039
VERSION I09039.1 GI:588252
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE
  1 (bases 1 to 952)
  AUTHORS Reddy,V.B.
  TITLE SUBSTANCES FOR DETECTING BLOODCLOTS
  JOURNAL Patent: WO 8808878-A 2 17-NOV-1988;
  FEATURES
    source
      Location/Qualifiers
        1..952
        /organism="unknown"
      225 a 266 c 286 g 175 t

BASE COUNT 225 a 266 c 286 g 175 t
ORIGIN
Query Match
  Best Local Similarity 100.0%; Score 18; DB 6; Length 952;
  Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCTGAGGGAACAGTGAC 18
    |||||

```



Db 638 TCTGAGGGAACAGTGAC 655

RESULT 6  
A27447  
LOCUS A27447 1065 bp DNA linear PAT 09-JUL-2002  
DEFINITION DNA sequence of coding region in pthrttpr.  
ACCESSION A27447  
VERSION A27447.1 GI:21727235  
KEYWORDS  
SOURCE synthetic construct  
ORGANISM synthetic construct  
artificial sequences.  
REFERENCE 1 (bases 1 to 1065)  
AUTHORS Niwa,M., Saito,Y., Sasaki,H., Hayashi,M., Notani,J. and Kobayashi,M.  
TITLE New tissue plasminogen activator  
JOURNAL Patent: EP 0302456-A 47 08-FEB-1989;  
FUJISAWA PHARMACEUTICAL CO., LTD  
FEATURES Location/Qualifiers  
source  
/organism="synthetic construct"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:32630"  
BASE COUNT 225 a 314 c 311 g 215 t  
ORIGIN

Query Match 100.0%; Score 18; DB 6; Length 1065;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCTGAGGGAACAGTGAC 18  
|||||  
Db 4 TCTGAGGGAACAGTGAC 21

RESULT 7  
AR059995  
LOCUS AR059995 1065 bp DNA linear PAT 29-SEP-1999  
DEFINITION Sequence 60 from patent US 5840533.  
ACCESSION AR059995  
VERSION AR059995.1 GI:5986445  
KEYWORDS Unknown.  
SOURCE Unknown.  
ORGANISM Unclassified.  
REFERENCE 1 (bases 1 to 1065)  
AUTHORS Niwa,M., Saito,Y., Sasaki,H., Hayashi,M., Notani,J. and Kobayashi,M.  
TITLE Tissue plasminogen activator  
JOURNAL Patent: US 5840533-A 60 24-NOV-1998;  
FEATURES Location/Qualifiers  
source  
1..1065  
/organism="unknown"  
BASE COUNT 225 a 314 c 311 g 215 t  
ORIGIN

Query Match 100.0%; Score 18; DB 6; Length 1065;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCTGAGGGAACAGTGAC 18  
|||||  
Db 4 TCTGAGGGAACAGTGAC 21

RESULT 8  
AX431270  
LOCUS AX431270 1065 bp DNA linear PAT 28-JUN-2002  
DEFINITION Sequence 4 from Patent WO240650.  
ACCESSION AX431270  
VERSION AX431270.1 GI:21656152  
KEYWORDS

SOURCE synthetic construct  
synthetic construct  
artificial sequences.  
REFERENCE 1  
AUTHORS Goetz,F., Werner,R.G., Tayapiwatana,C., Manosroi,A. and Manosroi,J.  
TITLE Methods for large scale production of recombinant dna-derived tpa or k2s molecules  
JOURNAL Patent: WO 0240650-A 4 23-MAY-2002;  
BOEHRINGER INGELHEIM INT (DE)  
FEATURES Location/Qualifiers  
source  
1..1065  
/organism="synthetic construct"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:32630"  
/note="coding sequence for K2S protein"  
BASE COUNT 225 a 314 c 312 g 214 t  
ORIGIN

Query Match 100.0%; Score 18; DB 6; Length 1065;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCTGAGGGAACAGTGAC 18  
|||||  
Db 1 TCTGAGGGAACAGTGAC 18

RESULT 10  
E01943  
LOCUS E01943 1065 bp DNA linear PAT 29-SEP-1997  
DEFINITION Synthetic DNA encoding new t-PA tissue plasminogen activator.  
ACCESSION E01943  
VERSION E01943.1 GI:2170191  
KEYWORDS JP 1989104167-A/10.  
SOURCE synthetic construct  
ORGANISM synthetic construct  
artificial sequences.  
REFERENCE 1 (bases 1 to 1065)  
AUTHORS Niwa,M., Saito,Y., Sasaki,H., Hayashi,M., Notani,J. and

Kobayashi,M.  
NOVEL TISSUE PLASMINOGEN ACTIVATION FACTOR  
Patent: JP 1989104167-A 10 21-APR-1989;  
FUJISAWA PHARMACEUT CO LTD  
OS Artificial gene  
OC Artificial sequence; Genes.  
PN JP 1989104167-A/10  
PD 21-APR-1989  
PF 01-AUG-1988 JP 1988192320  
PR 03-AUG-1987 GB 87 8718298, 26-OCT-1987 GB 87 8725052, PR  
13-NOV-1987 GB 87 8726683  
PI NIWA MINEO, SAITO YOSHIMASA, SASAKI HITOSHI, HAYASHI MASAKO,  
PI NOTANI JOJI,  
PI KOBAYASHI MASAKAZU  
PC C12N9/64, C12N1/20, C12N5/00, C12N15/00//A61K37/54, (C12N9/64, PC  
C12R1.19),  
PC (C12N9/64, C12R1.91);  
CC strandedness: Double;  
CC topology: Linear;  
CC hypothetical: No;  
CC anti-sense: No;  
CC \*source: Clone-pthTtrp; Location/Qualifiers  
FH Key  
FH CDS  
FT CDS  
FT Location/Qualifiers  
FT 1..1065  
/product='new t-PA'.  
/organism='synthetic construct'  
/mol\_type='genomic DNA'  
/db\_xref='taxon:32630'

BASE COUNT 225 a 314 c 311 g 215 t  
ORIGIN

Query Match 100.0%; Score 18; DB 6; Length 1065;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCTGAGGGAACAGTGAC 18  
|||||  
Db 4 TCTGAGGGAACAGTGAC 21

RESULT 11  
LOCUS A27431 1068 bp DNA linear PAT 09-JUL-2002  
DEFINITION DNA sequence of coding region in pTTPAdeltatrp.  
ACCESSION A27431  
VERSION A27431.1 GI:21727228  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
FEATURES  
source

1..1068  
/organism='synthetic construct'  
/mol\_type='genomic DNA'  
/db\_xref='taxon:32630'

BASE COUNT 226 a 314 c 313 g 215 t  
ORIGIN

Query Match 100.0%; Score 18; DB 6; Length 1068;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCTGAGGGAACAGTGAC 18  
|||||

Db 4 TCTGAGGGAACAGTGAC 21

RESULT 12  
LOCUS A27433

DEFINITION DNA sequence of coding region in pTTPAdeltatrp.  
ACCESSION A27433  
VERSION A27433.1 GI:21727229  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
FEATURES  
source

1..1068  
/organism='synthetic construct'  
/mol\_type='genomic DNA'  
/db\_xref='taxon:32630'

BASE COUNT 225 a 315 c 312 g 216 t  
ORIGIN

Query Match 100.0%; Score 18; DB 6; Length 1068;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCTGAGGGAACAGTGAC 18  
|||||  
Db 4 TCTGAGGGAACAGTGAC 21

RESULT 13  
LOCUS A27439

DEFINITION DNA sequence of coding region in pSTTKtrp.  
ACCESSION A27439  
VERSION A27439.1 GI:21727232  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
FEATURES  
source

1..1068  
/organism='synthetic construct'  
/mol\_type='genomic DNA'  
/db\_xref='taxon:32630'

BASE COUNT 228 a 313 c 312 g 215 t  
ORIGIN

Query Match 100.0%; Score 18; DB 6; Length 1068;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCTGAGGGAACAGTGAC 18  
|||||  
Db 4 TCTGAGGGAACAGTGAC 21

RESULT 14  
LOCUS A27445

DEFINITION DNA sequence of coding region in putTtrp.

ACCESSION A27445  
 VERSION A27445.1 GI:21727234  
 KEYWORDS  
 SOURCE synthetic construct  
 ORGANISM synthetic construct  
 artificial sequences  
 REFERENCE 1 (bases 1 to 1068)  
 AUTHORS Niwa,M., Saito,Y., Sasaki,H., Hayashi,M., Notani,J. and Kobayashi,M.  
 TITLE New tissue plasminogen activator  
 JOURNAL Patent: EP 0302456-A 45 08-FEB-1989;  
 FUJISAWA PHARMACEUTICAL CO., LTD  
 FEATURES Location/Qualifiers  
 source  
 1..1068  
 /organism="synthetic construct"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:32630"  
 BASE COUNT 226 a 314 c 313 g 215 t  
 ORIGIN

Query Match 100.0%; Score 18; DB 6; Length 1068;  
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCTGAGGGAACAGTGAC 18  
 ||||||||||||  
 Db 4 TCTGAGGGAACAGTGAC 21

RESULT 15  
 AR059987  
 LOCUS AR059987 1068 bp DNA linear PAT 29-SEP-1999  
 DEFINITION Sequence 44 from patent US 5840533.  
 ACCESSION AR059987  
 VERSION AR059987.1 GI:5986437  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unknown.  
 REFERENCE 1 (bases 1 to 1068)  
 AUTHORS Niwa,M., Saito,Y., Sasaki,H., Hayashi,M., Notani,J. and Kobayashi,M.  
 TITLE Tissue plasminogen activator  
 JOURNAL Patent: US 5840533-A 44 24-NOV-1998;  
 FEATURES Location/Qualifiers  
 source  
 1..1068  
 /organism="unknown"  
 BASE COUNT 226 a 314 c 313 g 215 t  
 ORIGIN

Query Match 100.0%; Score 18; DB 6; Length 1068;  
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCTGAGGGAACAGTGAC 18  
 ||||||||||||  
 Db 4 TCTGAGGGAACAGTGAC 21

Search completed: August 19, 2003, 04:46:24  
 Job time : 77.5794 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 18, 2003, 19:51:53 ; Search time 6.34127 Seconds  
(without alignments)  
7662.485 Million cell updates/sec

Title: US-09-987-455-1

Perfect score: 18

Sequence: 1 tctgagggaacagtgc 18

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N\_Geneseq 19Jun03;\*

- 1: /SIDSI/gcgdata/genseq/genseq-n-emb1/NA1980.DAT.\*
- 2: /SIDSI/gcgdata/genseq/genseq-n-emb1/NA1981.DAT.\*
- 3: /SIDSI/gcgdata/genseq/genseq-n-emb1/NA1982.DAT.\*
- 4: /SIDSI/gcgdata/genseq/genseq-n-emb1/NA1983.DAT.\*
- 5: /SIDSI/gcgdata/genseq/genseq-n-emb1/NA1984.DAT.\*
- 6: /SIDSI/gcgdata/genseq/genseq-n-emb1/NA1985.DAT.\*
- 7: /SIDSI/gcgdata/genseq/genseq-n-emb1/NA1986.DAT.\*
- 8: /SIDSI/gcgdata/genseq/genseq-n-emb1/NA1987.DAT.\*
- 9: /SIDSI/gcgdata/genseq/genseq-n-emb1/NA1988.DAT.\*
- 10: /SIDSI/gcgdata/genseq/genseq-n-emb1/NA1989.DAT.\*
- 11: /SIDSI/gcgdata/genseq/genseq-n-emb1/NA1990.DAT.\*
- 12: /SIDSI/gcgdata/genseq/genseq-n-emb1/NA1991.DAT.\*
- 13: /SIDSI/gcgdata/genseq/genseq-n-emb1/NA1992.DAT.\*
- 14: /SIDSI/gcgdata/genseq/genseq-n-emb1/NA1993.DAT.\*
- 15: /SIDSI/gcgdata/genseq/genseq-n-emb1/NA1994.DAT.\*
- 16: /SIDSI/gcgdata/genseq/genseq-n-emb1/NA1995.DAT.\*
- 17: /SIDSI/gcgdata/genseq/genseq-n-emb1/NA1996.DAT.\*
- 18: /SIDSI/gcgdata/genseq/genseq-n-emb1/NA1997.DAT.\*
- 19: /SIDSI/gcgdata/genseq/genseq-n-emb1/NA1998.DAT.\*
- 20: /SIDSI/gcgdata/genseq/genseq-n-emb1/NA1999.DAT.\*
- 21: /SIDSI/gcgdata/genseq/genseq-n-emb1/NA2000.DAT.\*
- 22: /SIDSI/gcgdata/genseq/genseq-n-emb1/NA2001A.DAT.\*
- 23: /SIDSI/gcgdata/genseq/genseq-n-emb1/NA2001B.DAT.\*
- 24: /SIDSI/gcgdata/genseq/genseq-n-emb1/NA2002.DAT.\*
- 25: /SIDSI/gcgdata/genseq/genseq-n-emb1/NA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	18	100.0	18	AAD40613	OmpA N-terminal pe
2	18	100.0	18	AAD38989	Human K2S heterolo
3	18	100.0	42	AAT15496	Megakaryocyte diff
4	18	100.0	42	AAD40621	Human tPA gene amp
5	18	100.0	42	AAD38990	Human tissue plas
6	18	100.0	55	AAT16011	Human tissue-plasm
7	18	100.0	952	9 AAN80896	Recombinant human
8	18	100.0	1047	14 AAO40658	tPA-2 cDNA. Synth

9	18	100.0	1047	14	AAQ40664	tPA-8 cDNA. Synth
10	18	100.0	1065	24	AAD40616	Human kringle 2 se
11	18	100.0	1066	10	AAN91128	Sequence of coding
12	18	100.0	1068	10	AAN91120	Sequence of coding
13	18	100.0	1068	10	AAN91122	Sequence of coding
14	18	100.0	1068	10	AAN91124	Sequence of coding
15	18	100.0	1068	14	AAQ53320	Human truncated tp
16	18	100.0	1068	14	AAQ53322	Human truncated tp
17	18	100.0	1068	14	AAQ40657	tPA-1 cDNA. Synth
18	18	100.0	1068	14	AAQ40660	tPA-6 cDNA. Synth
19	18	100.0	1068	14	AAQ40663	tPA-7 cDNA. Synth
20	18	100.0	1068	14	AAQ40667	tPA-11 cDNA. Synth
21	18	100.0	1068	14	AAQ40669	tPA-12 cDNA. Synth
22	18	100.0	1068	14	AAQ40671	tPA-13 cDNA. Synth
23	18	100.0	1068	14	AAQ40673	tPA-14 cDNA. Synth
24	18	100.0	1068	14	AAQ40677	tPA-16 cDNA. Synth
25	18	100.0	1068	14	AAQ40679	tPA-17 cDNA. Synth
26	18	100.0	1068	14	AAQ40681	tPA-18 cDNA. Synth
27	18	100.0	1068	14	AAQ40683	tPA-5 cDNA. Synth
28	18	100.0	1068	14	AAQ40685	tPA-6 cDNA. Synth
29	18	100.0	1068	14	AAQ40675	tPA-15 cDNA. Synth
30	18	100.0	1125	10	AAN91129	Sequence of coding
31	18	100.0	1128	24	AAD40614	OmpA-K2S fusion pr
32	18	100.0	1170	10	AAN91135	Sequence of coding
33	18	100.0	1170	10	AAN91136	Sequence of coding
34	18	100.0	1314	10	AAN91117	Sequence of coding
35	18	100.0	1314	10	AAN91123	Sequence of coding
36	18	100.0	1314	10	AAN91126	Sequence of coding
37	18	100.0	1314	10	AAN91127	Sequence of coding
38	18	100.0	1341	11	AAQ06782	Sequence encoding
39	18	100.0	1419	10	AAN91133	Sequence of coding
40	18	100.0	1437	9	AAN81069	Mutant no. 2600: t
41	18	100.0	1440	9	AAN81492	Modified tissue pl
42	18	100.0	1440	9	AAN82183	Modified tissue pl
43	18	100.0	1440	19	AAV47648	Fibroblast growth
44	18	100.0	1440	20	AAK32220	Human FGFR1-tPA fu
45	18	100.0	1440	20	AAK59330	cDNA coding for hu

ALIGNMENTS

RESULT 1

AAD40613	
ID	AAD40613 standard; DNA; 18 BP.
XX	
AC	AAD40613;
XX	
DT	30-OCT-2002 (first entry)
XX	
DE	OmpA N-terminal peptide encoding DNA #1.
XX	
KW	Tissue plasminogen activator; tPA; kringle 2 serine protease; stroke;
KW	K2S; cardiac infarction; acute myocardial infarction; pulmonary embolism;
KW	artery occlusion; deep vein thrombosis; blood clotting; thrombolytic;
KW	cerebroprotective; cardiact; ompA; gene; ds.
XX	
OS	Unidentified.
XX	
FH	Key
CDS	Location/Qualifiers
FT	1..18
FT	/*tag= a
FT	/product= "OmpA N-terminal peptide"
FT	/note= "No start and stop codon"
FT	/partial
XX	
XX	
XX	WO2002040650-A2.
PN	
XX	
PD	23-MAY-2002.
XX	
XX	
XX	07-NOV-2001; 2001WO-EPI2857.
XX	
XX	
PR	14-NOV-2000; 2000GB-0027779.

XX (BOEH ) BOEHRINGER INGELHEIM INT GMBH.  
 PA Werner R, Goetz F, Tayapiwatana C, Manosroi J, Manosroi A;  
 PI WPI; 2002-519376/55.  
 XX P-PSDB; AAE25033.  
 XX Producing active, correctly folded recombinant tissue plasminogen  
 PT activator, Kringle 2 serine protease in prokaryotic cells by expressing  
 PT the protein-encoding DNA operably linked to DNA coding for signal  
 XX peptide OmpA -  
 PS Claim 2; Page 29; 80pp; English.  
 XX The present invention relates to a method of producing extracellularly  
 CC secreted, active, correctly folded, recombinant tissue plasminogen  
 CC activator (tPA), Kringle 2 serine protease molecule (K2S) or their  
 CC variants in prokaryotic cells by expressing the protein-encoding DNA  
 CC operably linked to DNA coding for signal peptide OmpA. The method is  
 CC useful for producing recombinant DNA-derived tPA, K2S or their variants.  
 CC Sequences of the invention are useful for manufacturing a medicament  
 CC for treating stroke, cardiac infarction, acute myocardial infarction,  
 CC pulmonary embolism, any artery occlusion such as intracranial artery  
 CC occlusion (e.g., arteries supplying the brain), peripherally occluded  
 CC arteries, coronary artery occlusion, deep vein thrombosis or related  
 CC diseases associated with unwanted blood clotting. The present sequence  
 CC is a DNA fragment encoding OmpA N-terminal peptide.  
 XX  
 SQ Sequence 18 BP; 6 A; 3 C; 6 G; 3 T; 0 other;  
 Query Match 100.0%; Score 18; DB 24; Length 18;  
 Best Local Similarity 100.0%; Pred. No. 10;  
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 TCTGAGGGAACACGTGAC 18  
 DB 1 TCTGAGGGAACACGTGAC 18  
 RESULT 2  
 AAD38989  
 ID AAD38989 standard; DNA; 18 BP.  
 AC AAD38989;  
 XX  
 XX 23-SEP-2002 (first entry)  
 DT Human K2S heterologous protein N-terminal peptide DNA #2.  
 DE Recombinant DNA-derived heterologous protein; OmpA; K2S; kringle 2;  
 XX serine protease; human; ds.  
 KW Homo sapiens.  
 XX  
 XX Key Location/Qualifiers  
 FH 1..12  
 CDS /\*tag= a  
 FT /product= "Human K2S heterologous protein N-terminal  
 FT peptide #2".  
 XX WO200240696-A2.  
 XX  
 XX 23-MAY-2002.  
 PD  
 XX 08-NOV-2001; 2001WO-EP12920.  
 PF  
 XX 14-NOV-2000; 2000GB-0027782.  
 PR  
 XX (BOEH ) BOEHRINGER INGELHEIM INT GMBH.  
 PA Werner R, Goetz F, Tayapiwatana C, Manosroi J, Manosroi A;  
 XX PI

DR WPI; 2002-471625/50.  
 XX Producing recombinant DNA-derived kringle 2 plus serine protease.  
 PT comprises using a prokaryotic cell expressing a vector having a DNA  
 PT coding for a heterologous protein operably linked to a DNA coding for  
 PT the signal peptide OmpA -  
 XX  
 XX Disclosure; Page 22; 52pp; English.  
 PS  
 XX The invention relates to a method for producing recombinant DNA-derived  
 CC heterologous protein in prokaryotic cells, where the heterologous protein  
 CC is secreted extracellularly as an active and correctly folded protein and  
 CC the prokaryotic cell contains and expresses a vector comprising the DNA  
 CC coding for the heterologous protein operably linked to the DNA coding  
 CC for the signal peptide OmpA or its functional derivative. The method is  
 CC useful for commercial large-scale production of heterologous proteins;  
 CC e.g. K2S (kringle 2 plus serine protease), in prokaryotic cells, and is  
 CC generally applicable in the expression of several different proteins and  
 CC polypeptides which do not require mammalian glycosylation in prokaryotic  
 CC host cells. The method may also be used to obtain DNA sequences of a  
 CC protein of interest to be expressed from databases and cloned for use.  
 CC The present sequence is human K2S heterologous protein N-terminal peptide  
 CC DNA.  
 XX  
 SQ Sequence 18 BP; 6 A; 3 C; 6 G; 3 T; 0 other;  
 Query Match 100.0%; Score 18; DB 24; Length 18;  
 Best Local Similarity 100.0%; Pred. No. 10;  
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 TCTGAGGGAACACGTGAC 18  
 DB 1 TCTGAGGGAACACGTGAC 18  
 RESULT 3  
 AAT15496/c  
 ID AAT15496 standard; DNA; 42 BP.  
 XX  
 XX AAT15496;  
 AC  
 XX 20-SEP-1996 (first entry)  
 DT  
 XX Megakaryocyte differentiation/proliferation factor, NTPO cDNA primer.  
 DE  
 XX Megakaryocyte; differentiation factor; proliferation; recombinant;  
 KW thrombocytopenia; NTPO; polymerase chain reaction; ss.  
 XX Synthetic.  
 XX OS  
 XX WO9603434-A1.  
 PN  
 XX 08-FEB-1996.  
 PD  
 XX 25-JUL-1995; 95WO-JP01476.  
 PF  
 XX 25-JUL-1994; 94JP-0172497.  
 PR  
 XX (EISA ) EISAI CO LTD.  
 PA  
 XX Komatsu N, Okada M, Yoshitake S;  
 PI  
 XX WPI; 1996-117001/12.  
 DR  
 XX Megakaryocyte differentiation/proliferation factor - and fragments  
 PT of it are prepared by recombinant techniques and used for treatment  
 PT of thrombocytopenia  
 XX  
 XX Example 1; Page 12; 41pp; Japanese.  
 PS  
 XX AAT15487-T15498 are PCR primers used for the isolation and amplification  
 CC of cDNA encoding a human megakaryocyte differentiation/proliferation  
 CC factor, NTPO, some of the primers may also be used to engineer  
 CC

CC recombinant NTPO DNA. DNA encoding the factor can be used to produce  
 CC a recombinant NTPO at a high yield. NTPO is useful for the treatment  
 CC of thrombocytopenia.

SQ Sequence 42 BP; 10 A; 9 C; 11 G; 12 T; 0 other;

Query Match 100.0%; Score 18; DB 17; Length 42;

Best Local Similarity 100.0%; Pred. No. 11;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 TCTGAGGGAACAGTGAC 18

Db 31 TCTGAGGGAACAGTGAC 14

RESULT 4

AAD40621

ID AAD40621 standard; DNA; 42 BP.

XX

AC AAD40621;

XX 30-OCT-2002 (first entry)

XX Human tPA gene amplifying PCR primer, sk2/174.

XX Tissue plasminogen activator; tPA; kringle 2 serine protease; stroke;

KW K2S; cardiac infarction; acute myocardial infarction; pulmonary embolism;

KW artery occlusion; deep vein thrombosis; blood clotting; thrombolytic;

KW cerebroprotective; cardiant; PCR; primer; human; ss.

XX Homo sapiens.

OS

XX WO200240650-A2.

XX 23-MAY-2002.

XX 07-NOV-2001; 2001WO-EPI2857.

XX 14-NOV-2000; 2000GB-0027779.

XX (BOEH) BOEHRINGER INGELHEIM INT GMBH.

XX Werner R, Goetz F, Tayapiwatana C, Manosroi J, Manosroi A;

XX WPI; 2002-519376/55.

XX Producing active, correctly folded recombinant tissue plasminogen  
 PT activator, Kringle 2 serine protease in prokaryotic cells by expressing  
 PT the protein-encoding DNA operably linked to DNA coding for signal  
 PT peptide OmpA -

XX Example 1; Page 80; 80pp; English.

XX The present invention relates to a method of producing extracellularly  
 CC secreted, active, correctly folded, recombinant tissue plasminogen  
 CC activator (tPA), Kringle 2 serine protease molecule (K2S) or their  
 CC variants in prokaryotic cells by expressing the protein-encoding DNA  
 CC operably linked to DNA coding for signal peptide OmpA. The method is  
 CC useful for producing recombinant DNA-derived tPA, K2S or their variants.  
 CC Sequences of the invention are useful for manufacturing a medicament  
 CC for treating stroke, cardiac infarction, acute myocardial infarction,  
 CC pulmonary embolism, any artery occlusion such as intracranial artery  
 CC occlusion (e.g., arteries supplying the brain), peripherally occluded  
 CC arteries, coronary artery occlusion, deep vein thrombosis or related  
 CC diseases associated with unwanted blood clotting. The present sequence  
 CC is a PCR primer which is used for amplifying human tissue plasminogen  
 CC activator (tPA) gene. This primer is used in the exemplification of  
 CC the invention.

XX Sequence 42 BP; 10 A; 9 C; 19 G; 4 T; 0 other;

Query Match 100.0%; Score 18; DB 24; Length 42;

Best Local Similarity 100.0%; Pred. No. 11;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 TCTGAGGGAACAGTGAC 18

Db 25 TCTGAGGGAACAGTGAC 42

RESULT 5

AAD38990

ID AAD38990 standard; DNA; 42 BP.

XX

AC AAD38990;

XX 23-SEP-2002 (first entry)

XX Human tissue plasminogen activator (tPA) gene amplifying primer, sk2/174.

XX Recombinant DNA-derived heterologous protein; OmpA; K2S; kringle 2;

KW serine protease; human; primer; PCR; ss.

XX Homo sapiens.

OS

XX WO200240696-A2.

XX 23-MAY-2002.

XX 08-NOV-2001; 2001WO-EPI2920.

XX 14-NOV-2000; 2000GB-0027782.

XX (BOEH) BOEHRINGER INGELHEIM INT GMBH.

XX Werner R, Goetz F, Tayapiwatana C, Manosroi J, Manosroi A;

XX WPI; 2002-471625/50.

XX Producing recombinant DNA-derived kringle 2 plus serine protease,  
 PT comprises using a prokaryotic cell expressing a vector having a DNA  
 PT coding for a heterologous protein operably linked to a DNA coding for  
 PT the signal peptide OmpA -

XX Example 1; Page 9; 52pp; English.

XX The invention relates to a method for producing recombinant DNA-derived  
 CC heterologous protein in prokaryotic cells, where the heterologous protein  
 CC is secreted extracellularly as an active and correctly folded protein and  
 CC the prokaryotic cell contains and expresses a vector comprising the DNA  
 CC coding for the heterologous protein operably linked to the DNA coding  
 CC for the signal peptide OmpA or its functional derivative. The method is  
 CC useful for commercial large-scale production of heterologous proteins,  
 CC e.g. K2S (kringle 2 plus serine protease), in prokaryotic cells, and is  
 CC generally applicable in the expression of several different proteins and  
 CC polypeptides which do not require mammalian glycosylation in prokaryotic  
 CC host cells. The method may also be used to obtain DNA sequences of a  
 CC protein of interest to be expressed from databases and cloned for use.  
 CC The present sequence is human tissue plasminogen activator (tPA) gene  
 CC amplifying PCR primer.

XX Sequence 42 BP; 10 A; 9 C; 19 G; 4 T; 0 other;

Query Match 100.0%; Score 18; DB 24; Length 42;

Best Local Similarity 100.0%; Pred. No. 11;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 TCTGAGGGAACAGTGAC 18

Db 25 TCTGAGGGAACAGTGAC 42

RESULT 6

AAT16011/c

ID AAT16011 standard; DNA; 55 BP.

XX

```

AC AAT16011;
XX
XX 23-SEP-1996 (first entry)
XX
XX Human tissue-plasminogen activator (t-PA) gene N-terminal PCR primer.
XX
XX Megakaryocyte; differentiation factor; proliferation; recombinant;
XX thrombocytopenia; NTPO; polymerase chain reaction; ss.
XX
XX Synthetic.
XX
XX WO9603434-A1.
XX
XX 08-FEB-1996.
XX
XX 25-JUL-1995; 95WO-JP01476.
XX
XX 25-JUL-1994; 94JP-0172497.
XX
XX (EISA) EISAI CO LTD.
XX
XX Komatsu N, Okada M, Yoshitake S;
XX
XX WPI; 1996-117001/12.
XX
XX Megakaryocyte differentiation/proliferation factor - and fragments
XX of it are prepared by recombinant techniques and used for treatment
XX of thrombocytopenia
XX
XX Example 4; Page 13; 41pp; Japanese.
XX
XX AAT16011-T16014 are PCR primers used for the isolation and amplification
XX of the human tissue-plasminogen activator (t-PA) gene. The t-PA gene
XX may be incorporated into a plasmid contg. DNA encoding a megakaryocyte
XX differentiation/proliferation factor, NTPO, so as expression produces
XX a chimeric t-PA-NTPO protein. DNA encoding the factor can be used to
XX produce a recombinant NTPO at a high yield. NTPO is useful for the
XX treatment of thrombocytopenia either alone or in conjunction with
XX t-PA.
XX
XX Sequence 55 BP; 12 A; 11 C; 17 G; 15 T; 0 other;
XX
XX Query Match 100.0%; Score 18; DB 17; Length 55;
XX Best Local Similarity 100.0%; Pred. No. 12;
XX Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCTGAGGGAACAGTGAC 18
Db 40 TCTGAGGGAACAGTGAC 23

RESULT 7
AAN80896
ID AAN80896 standard; cDNA; 952 BP.
XX
XX AAN80896;
XX
XX 31-OCT-1990 (first entry)
XX
XX Recombinant human tissue plasminogen activator A chain sequence.
XX
XX human tissue plasminogen activator (t-PA) A chain;
XX antithrombotic agents; ds.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX sig_peptide 14..118
XX FT /*tag= a
XX FT mat_peptide 119..943
XX FT /*tag= b
XX FT /product=t-PA A-chain
XX FT misc_feature 915..952

```

```

FT /*tag= c
FT /label=synthetic oligomer
FT /note="capable of being transcribed into an mRNA
FT stop codon"
XX
XX WO8808878-A.
XX
XX 17-NOV-1988.
XX
XX 12-MAY-1988; 88WO-US01624.
XX
XX 15-MAY-1987; 87US-0050950.
XX
XX 12-MAY-1988; 88WO-US01624.
XX
XX (INTE-) INTEG GENETICS INC.
XX
XX Reddy VB;
XX
XX WPI; 1988-338227/47.
XX
XX P-PSDB; AAP80770.
XX
XX DNA coding for tissue plasminogen activator A chain -
XX used as labelled A chain for antithrombotic drug screening
XX
XX Disclosure; Page ?; p; English.
XX
XX Recombinant sequence was obtained upon ligation of a 184bp
XX DdeI-EcoRI fragment of p7BDI-10, a 730bp Sall-EcoRI fragment of
XX p7BDT-10 encoding N-terminal portion of t-PA and a synthetic
XX oligomer capable of being transcribed into an mRNA stop codon.
XX The ligated product was cleaved with Sall to give the 952bp
XX fragment shown. This fragment was cloned into the XhoI site of
XX bovine papilloma virus DNA and used to transfect C127 mouse cells.
XX
XX Sequence 952 BP; 225 A; 266 C; 286 G; 175 T; 0 other;
XX
XX Query Match 100.0%; Score 18; DB 9; Length 952;
XX Best Local Similarity 100.0%; Pred. No. 16;
XX Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCTGAGGGAACAGTGAC 18
Db 638 TCTGAGGGAACAGTGAC 655

RESULT 8
AAQ40658
ID AAQ40658 standard; cDNA; 1047 BP.
XX
XX AAQ40658;
XX
XX 06-AUG-1993 (first entry)
XX
XX tPA-2 cDNA.
XX
XX Blood; tissue plasminogen activator; tPA; mutein; stability; tPA-1;
XX physiological; activity; tPA-2; PTB 1127; mutagenesis; plasmid;
XX truncated; deletion; ss.
XX
XX Synthetic.
XX
XX Key Location/Qualifiers
XX mutation 369..370
XX FT /*tag= a
XX FT /note= "position of deletion mutation"
XX
XX JP05076361-A.
XX
XX 30-MAR-1993.
XX
XX 10-MAY-1991; 91JP-0105689.
XX
XX 10-MAY-1990; 90JP-0118710.

```



PR 25-DEC-1990; 90JP-0405848.  
XX (TAKE ) TAKEDA CHEM IND LTD.  
XX WPI; 1993-139567/17.  
DR P-PSDB; AAR35401.  
XX Tissue plasminogen activator mutein - useful for treating  
PT myocardial infarction and cerebral thrombosis  
XX Disclosure; Page 26-27; 92pp; Japanese.  
PS  
XX The sequences given in AAQ0657-58 are cDNAs encoding tissue plasminogen  
CC activator (tPA) muteins, tPA-1 and tPA-2. These muteins are truncated  
CC versions of tPA (see AAQ0654) and tPA-2 further contains a deletion of  
CC amino acids corresponding to amino acids 296-302 of tPA. The plasmid  
CC PTB 1127 was treated with the synthetic oligomers given in AAQ0655-56  
CC in an in vitro mutagenesis system. The tPA muteins, tPA-1 and -2 have  
CC good stability in blood and good physiological activity.  
XX  
SQ Sequence 1047 BP; 220 A; 307 C; 305 G; 215 T; 0 other;  
Query Match 100.0%; Score 18; DB 14; Length 1047;  
Best Local Similarity 100.0%; Pred. No. 16;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 TCTGAGGGAACAGTGAC 18  
Db 4 TCTGAGGGAACAGTGAC 21  
RESULT 9  
ID AAQ0664 standard; cDNA; 1047 BP.  
AC AAQ0664;  
XX  
DT 06-AUG-1993 (first entry)  
DE tPA-8 cDNA.  
XX  
KW Blood; tissue plasminogen activator; tPA; mutein; stability; tPA-7;  
KW physiological; activity; tPA-8; PTB 1277; mutagenesis; plasmid;  
KW truncated; deletion; mutation; ss.  
XX  
OS Synthetic.  
XX  
FH Key Location/Qualifiers  
FT mutation 369..370  
FT /\*tag= a  
FT /note= "Position of deletion mutation"  
FT mutation 371..376  
FT /\*tag= b  
FT /note= "GAGCGG>TATCAC"  
XX  
PN JP05076361-A.  
XX  
PD 30-MAR-1993.  
XX  
PF 10-MAY-1991; 91JP-0105689.  
XX  
PR 10-MAY-1990; 90JP-0118710.  
PR 25-DEC-1990; 90JP-0405848.  
XX  
PA (TAKE ) TAKEDA CHEM IND LTD.  
XX  
XX WPI; 1993-139567/17.  
DR P-PSDB; AAR35404.  
XX Tissue plasminogen activator mutein - useful for treating  
PT myocardial infarction and cerebral thrombosis  
XX Disclosure; Page 31-32; 92pp; Japanese.

XX The sequences given in AAQ0663-64 are cDNAs encoding tissue plasminogen  
CC activator (tPA) muteins, tPA-7 and tPA-8. These muteins are truncated  
CC versions of tPA (see AAQ0654). tPA-7 further comprises a Glu-Arg>  
CC Tyr-His mutation amino acids corresponding to amino acids 303-304 of  
CC tPA and tPA-8 further contains a deletion of amino acids corresponding  
CC to amino acids 296-302 of tPA, and the mutation contained in tPA-7.  
CC The plasmid PTB 1277 was treated with the synthetic oligomers given  
CC in AAQ0661-62 in an in vitro mutagenesis system. The tPA muteins,  
CC tPA-7 and -8 have good stability in blood and good physiological  
CC activity.  
XX  
SQ Sequence 1047 BP; 221 A; 308 C; 301 G; 217 T; 0 other;  
Query Match 100.0%; Score 18; DB 14; Length 1047;  
Best Local Similarity 100.0%; Pred. No. 16;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 TCTGAGGGAACAGTGAC 18  
Db 4 TCTGAGGGAACAGTGAC 21  
RESULT 10  
ID AAD40616 standard; DNA; 1065 BP.  
AC AAD40616;  
XX  
DT 30-OCT-2002 (first entry)  
DE Human kringle 2 serine protease (K2S) DNA.  
XX  
KW Tissue plasminogen activator; tPA; kringle 2 serine protease; stroke;  
KW K2S; cardiac infarction; acute myocardial infarction; pulmonary embolism;  
KW artery occlusion; deep vein thrombosis; blood clotting; thrombolytic;  
KW cerebroprotective; cardiant; ompa; enzyme; human; gene; ds.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT CDS 1..1065  
FT /\*tag= a  
FT /product= "Human K2S protein"  
XX  
PN WO2002040650-A2.  
XX  
PD 23-MAY-2002.  
XX  
PF 07-NOV-2001; 2001WO-EPI2857.  
XX  
PR 14-NOV-2000; 2000GB-0027779.  
XX  
PA (BOEH ) BOEHRINGER INGELHEIM INT GMBH.  
XX  
XX Werner R, Goetz F, Tayapiwatana C, Manosroi J, Manosroi A;  
PI WPI; 2002-519376/55.  
DR P-PSDB; AAE25036.  
XX  
XX Producing active, correctly folded recombinant tissue plasminogen  
PT activator, Kringle 2 serine protease in prokaryotic cells by expressing  
PT the protein-encoding DNA operably linked to DNA coding for signal  
PT peptide Ompa  
XX  
PS Claim 12; Page 31; 80pp; English.  
XX  
XX The present invention relates to a method of producing extracellularly  
CC secreted, active, correctly folded, recombinant tissue plasminogen  
CC activator (tPA), Kringle 2 serine protease molecule (K2S) or their  
CC variants in prokaryotic cells by expressing the protein-encoding DNA  
CC operably linked to DNA coding for signal peptide Ompa. The method is  
CC useful for producing recombinant DNA-derived tPA, K2S or their variants.

CC Sequences of the invention are useful for manufacturing a medicament  
 CC for treating stroke, cardiac infarction, acute myocardial infarction,  
 CC pulmonary embolism, any artery occlusion such as intracranial artery  
 CC occlusion (e.g., arteries supplying the brain), peripherally occluded  
 CC arteries, coronary artery occlusion, deep vein thrombosis or related  
 CC diseases associated with unwanted blood clotting. The present sequence  
 CC is human K2S DNA.

XX SQ Sequence 1065 BP; 225 A; 314 C; 312 G; 214 T; 0 other;

Query Match 100.0%; Score 18; DB 24; Length 1065;  
 Best Local Similarity 100.0%; Pred. No. 17;  
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 TCTGAGGGAACAGTGAC 18

Db 1 TCTGAGGGAACAGTGAC 18

# RESULT 11

AA91128  
 ID AA91128 standard; DNA; 1066 BP.

XX AC AA91128;

XX 25-MAR-2003 (updated)  
 DT 03-OCT-2002 (updated)  
 DT 18-JUN-1990 (first entry)

XX Sequence of coding region in plasmid pUTtrp.

XX Tissue plasminogen activator; tPA; thrombolytic agent;

XX plasminogen; vascular diseases.

XX Synthetic.

XX FH Key Location/Qualifiers  
 FT CDS 1..1065  
 FT /\*tag= a

XX EP302456-A.

XX 08-FEB-1989.

XX 02-AUG-1988; 88EP-0112569.

XX 03-AUG-1987; 87GB-0018298.

XX 26-OCT-1987; 87GB-0025052.

XX 13-NOV-1987; 87GB-0026683.

XX (FUJI ) FUJISAWA PHARM CO LTD.

XX Niwa M, Saito Y, Sasaki H, Hayashi M, Notani J, Kobayashi M;

XX WPI; 1989-040625/06.

XX P-PSDB; AAP94414.

XX New tissue plasminogen activator -  
 PT comprising finger and growth factor domains lacking tPA for  
 PT longer half-life and stronger thrombolytic activity.

XX Disclosure; Page ?; 68pp; English.

XX (Updated on 03-OCT-2002 to add missing OS field.)  
 CC (Updated on 25-MAR-2003 to correct PA field.)

XX SQ Sequence 1066 BP; 225 A; 314 C; 312 G; 215 T; 0 other;

Query Match 100.0%; Score 18; DB 10; Length 1066;

Best Local Similarity 100.0%; Pred. No. 17;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 TCTGAGGGAACAGTGAC 18

Db 4 TCTGAGGGAACAGTGAC 21

# RESULT 12

AA91120  
 ID AA91120 standard; DNA; 1068 BP.

XX AC AA91120;

XX 25-MAR-2003 (updated)  
 DT 03-OCT-2002 (updated)  
 DT 18-JUN-1990 (first entry)

XX Sequence of coding region in plasmid pST112.

XX Tissue plasminogen activator; tPA; thrombolytic agent;

XX plasminogen; vascular diseases.

XX Synthetic.

XX FH Key Location/Qualifiers  
 FT CDS 1..1065  
 FT /\*tag= a

XX EP302456-A.

XX 08-FEB-1989.

XX 02-AUG-1988; 88EP-0112569.

XX 03-AUG-1987; 87GB-0018298.

XX 26-OCT-1987; 87GB-0025052.

XX 13-NOV-1987; 87GB-0026683.

XX (FUJI ) FUJISAWA PHARM CO LTD.

XX Niwa M, Saito Y, Sasaki H, Hayashi M, Notani J, Kobayashi M;

XX WPI; 1989-040625/06.

XX P-PSDB; AAP94407.

XX New tissue plasminogen activator -  
 PT comprising finger and growth factor domains lacking tPA for  
 PT longer half-life and stronger thrombolytic activity.

XX Disclosure; Page ?; 68pp; English.

XX (Updated on 03-OCT-2002 to add missing OS field.)

XX (Updated on 25-MAR-2003 to correct PA field.)

XX SQ Sequence 1068 BP; 226 A; 314 C; 313 G; 215 T; 0 other;

Query Match 100.0%; Score 18; DB 10; Length 1068;  
 Best Local Similarity 100.0%; Pred. No. 17;  
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 TCTGAGGGAACAGTGAC 18

Db 4 TCTGAGGGAACAGTGAC 21

# RESULT 13

AA91122  
 ID AA91122 standard; DNA; 1068 BP.

XX AC AA91122;

XX 25-MAR-2003 (updated)  
 DT 03-OCT-2002 (updated)  
 DT 18-JUN-1990 (first entry)

XX Sequence of coding region in plasmid pTTiPA delta trp.

XX Tissue plasminogen activator; tPA; thrombolytic agent;  
 KW plasminogen; vascular diseases.  
 XX OS Synthetic.

XX Key Location/Qualifiers  
 FT CDS 1..1065  
 FT /\*tag= a

XX EP302456-A.

XX PD 08-FEB-1989.

XX PF 02-AUG-1988; 88EP-0112569.

XX PR 03-AUG-1987; 87GB-0018298.

XX PR 26-OCT-1987; 87GB-0025052.

XX PR 13-NOV-1987; 87GB-0026683.

XX PA (FUJI ) FUJISAWA PHARM CO LTD.

XX PI Niwa M, Saito Y, Sasaki H, Hayashi M, Notani J, Kobayashi M;

XX DR WPI; 1989-040625/06.

XX DR P-PSDB; AAP94408.

XX PT New tissue plasminogen activator -  
 PT comprising finger and growth factor domains lacking tPA for  
 PT longer half-life and stronger thrombolytic activity.

XX PS Disclosure; Page ?; 68pp; English.

XX CC (Updated on 03-OCT-2002 to add missing OS field.)

XX CC (Updated on 25-MAR-2003 to correct PA field.)

XX SQ Sequence 1068 BP; 225 A; 315 C; 312 G; 216 T; 0 other;

Query Match 100.0%; Score 18; DB 10; Length 1068;  
 Best Local Similarity 100.0%; Pred. No. 17;  
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCTGAGGGAACAGTGAC 18

Db 4 TCTGAGGGAACAGTGAC 21

RESULT 14

AA91124

ID AA91124 standard; DNA; 1068 BP.

XX AC AA91124;

XX DT 25-MAR-2003 (updated)

XX DT 03-OCT-2002 (updated)

XX DT 18-JUN-1990 (first entry)

XX DE Sequence of coding region in plasmid pSTKtrp.

XX KW Tissue plasminogen activator; tPA; thrombolytic agent;  
 KW plasminogen; vascular diseases.

XX OS Synthetic.

XX Key Location/Qualifiers  
 FT CDS 1..1065  
 FT /\*tag= a

XX EP302456-A.

XX PD 08-FEB-1989.

XX PF 02-AUG-1988; 88EP-0112569.

XX 03-AUG-1987; 87GB-0018298.  
 PR 26-OCT-1987; 87GB-0025052.  
 PR 13-NOV-1987; 87GB-0026683.  
 XX (FUJI ) FUJISAWA PHARM CO LTD.

XX Niwa M, Saito Y, Sasaki H, Hayashi M, Notani J, Kobayashi M;  
 WPI; 1989-040625/06.  
 DR P-PSDB; AAP94411.

XX PT New tissue plasminogen activator -  
 PT comprising finger and growth factor domains lacking tPA for  
 PT longer half-life and stronger thrombolytic activity.

XX PS Disclosure; Page ?; 68pp; English.

XX CC (Updated on 03-OCT-2002 to add missing OS field.)

XX CC (Updated on 25-MAR-2003 to correct PA field.)

XX SQ Sequence 1068 BP; 228 A; 313 C; 312 G; 215 T; 0 other;

Query Match 100.0%; Score 18; DB 10; Length 1068;  
 Best Local Similarity 100.0%; Pred. No. 17;  
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCTGAGGGAACAGTGAC 18

Db 4 TCTGAGGGAACAGTGAC 21

RESULT 15

AAQ5320

ID AAQ5320 standard; cDNA to mRNA; 1068 BP.

XX AC AAQ5320;

XX DT 20-JUN-1994 (first entry)

XX DE Human truncated tPA-1 mutein coding sequence.

XX KW Tissue plasminogen activator; t-PA; mutein; fibrin; antigen;  
 KW anti-fibrin; monoclonal antibody; hybridoma; thrombolysis  
 KW targeted antithrombotic agent; hybridisation probe;  
 KW Kringle domain; protease; ds.

XX OS Homo sapiens.

XX Key Location/Qualifiers  
 FT CDS 1..1068  
 FT /\*tag= a  
 FT /product= tPA-1  
 FT /note= "encodes Kringle-2 and protease domains"

XX PN JF05304992-A.

XX PD 19-NOV-1993.

XX PF 17-JUN-1992; 92JP-0158301.

XX PR 20-JUN-1991; 91JP-0148936.

XX PA (TAKE ) TAKEDA CHEM IND LTD.

XX WPI; 1993-408334/51.

XX DR P-PSDB; AAR44836.

XX PT Hybrid monoclonal antibody - used for prepn. of thrombolytic drug  
 PT having increased thrombolytic activity and specificity and  
 PT reduced reactivity to fibrinogen

XX PS Disclosure; Fig 7; 38pp; Japanese.

XX Site-directed mutagenesis was carried out on a human tPA cDNA  
CC template. The resulting mutated coding sequence encoded a Met  
CC residue followed by amino acids 174-527 (i.e. the Kringle 2  
CC and protease domains) of mature wild-type tPA. The mutein has  
CC increased thrombolytic activity and very little reactivity to  
CC fibrinogen. Monoclonal antibodies specific for fibrin are used in  
CC the production of bispecific monoclonal antibodies which also  
CC recognise truncated tPA muteins such as tPA-1 lacking the finger,  
CC EGF and Kringle 1 domains.

XX  
SQ Sequence 1068 BP; 226 A; 314 C; 313 G; 215 T; 0 other;  
Query Match 100.0%; Score 18; DB 14; Length 1068;  
Best Local Similarity 100.0%; Pred. No. 17;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCTGAGGGAACAGTGAC 18  
|||  
Db 4 TCTGAGGGAACAGTGAC 21  
|||

Search completed: August 18, 2003, 23:28:45  
Job time : 6.34127 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 18, 2003, 22:51:23 ; Search time 1.40873 Seconds  
(without alignments)  
5639.757 Million cell updates/sec

Title: US-09-987-455-1

Perfect score: 18  
Sequence: 1 tctgagggaacagtgc 18

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA.\*

1: /cgn2\_6/ptodata/2/ina/5A\_COMB.seq:\*  
2: /cgn2\_6/ptodata/2/ina/5B\_COMB.seq:\*  
3: /cgn2\_6/ptodata/2/ina/6A\_COMB.seq:\*  
4: /cgn2\_6/ptodata/2/ina/6B\_COMB.seq:\*  
5: /cgn2\_6/ptodata/2/ina/PCTUS\_COMB.seq:\*  
6: /cgn2\_6/ptodata/2/ina/backfiles.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	18	100.0	1065	2	US-08-811-949-60
2	18	100.0	1068	2	US-08-811-949-44
3	18	100.0	1068	2	US-08-811-949-46
4	18	100.0	1068	2	US-08-811-949-52
5	18	100.0	1068	2	US-08-811-949-58
6	18	100.0	1170	2	US-08-811-949-64
7	18	100.0	1170	2	US-08-811-949-66
8	18	100.0	1314	2	US-08-811-949-48
9	18	100.0	1314	2	US-08-811-949-50
10	18	100.0	1314	2	US-08-811-949-54
11	18	100.0	1314	2	US-08-811-949-56
12	18	100.0	1419	2	US-08-811-949-62
13	18	100.0	1724	6	5200340-5
14	18	100.0	1738	6	5200340-1
15	18	100.0	1848	3	US-08-814-412-10
16	18	100.0	1955	2	US-08-883-795A-39
17	18	100.0	1974	2	US-08-811-949-38
18	18	100.0	2101	2	US-08-811-949-42
19	18	100.0	2162	1	US-08-119-512-3
20	18	100.0	2162	1	US-08-488-015B-3
21	18	100.0	2162	1	US-08-488-015B-25
22	18	100.0	2457	6	534773-1
23	18	100.0	2497	6	5185259-2
24	18	100.0	7360	1	US-08-286-740-1
25	18	100.0	7360	5	PCT-US95-09576-1
26	15.4	85.6	1459	3	US-09-020-956-174
27	15.4	85.6	1459	3	US-09-030-607-174

28	15.4	85.6	1459	4	US-09-439-313-174	Sequence 174, App
29	15.4	85.6	1459	4	US-09-352-616A-174	Sequence 174, App
30	15.4	85.6	1459	4	US-09-232-149A-174	Sequence 174, App
31	15.4	85.6	3772	4	US-09-996-243-61	Sequence 61, Appli
c 32	15	83.3	5159	2	US-08-146-930-3	Sequence 3, Appli
c 33	15	83.3	5159	3	US-08-458-240-3	Sequence 3, Appli
c 34	15	83.3	5159	5	PCT-US93-03993-3	Sequence 3, Appli
c 35	15	83.3	35100	1	US-08-306-691B-19	Sequence 19, Appli
c 36	15	83.3	35100	5	PCT-US93-06251-19	Sequence 19, Appli
37	14.8	82.2	20	3	US-08-930-601-6	Sequence 6, Appli
38	14.8	82.2	3222	4	US-09-206-942-48	Sequence 48, Appli
39	14.8	82.2	3240	4	US-09-206-942-46	Sequence 46, Appli
40	14.8	82.2	64467	4	US-09-803-671B-3	Sequence 3, Appli
c 41	14.4	80.0	1881	5	PCT-US94-09752-2	Sequence 2, Appli
42	14.4	80.0	6908	4	US-09-620-312D-977	Sequence 977, App
43	14	77.8	34185	4	US-09-545-481-3	Sequence 3, Appli
c 44	13.8	76.7	89	2	US-08-373-190-51	Sequence 51, Appli
c 45	13.8	76.7	89	2	US-08-438-190A-51	Sequence 51, Appli

#### ALIGNMENTS

RESULT 1  
US-08-811-949-60  
; Sequence 60, Application US/08811949  
; Patent No. 5840533  
; GENERAL INFORMATION:  
; APPLICANT: NIWA, MINEO  
; APPLICANT: SAITO, YOSHIMASA  
; APPLICANT: SASAKI, HITOSHI  
; APPLICANT: HAYASHI, MASAKO  
; APPLICANT: NOTANI, JOUJI  
; APPLICANT: KOBAYASHI, MASAKAZU  
; TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR  
; NUMBER OF SEQUENCES: 67  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,  
; ADDRESSEE: P.C.  
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400  
; CITY: ARLINGTON  
; STATE: VA  
; COUNTRY: USA  
; ZIP: 22202  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/811,949  
; FILING DATE: 05-MAR-1997  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: OBLON, NORMAN F.  
; REGISTRATION NUMBER: 24,618  
; REFERENCE/DOCKET NUMBER: 18-966-0  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 703-413-3000  
; TELEFAX: 703-413-2220  
; INFORMATION FOR SEQ ID NO: 60:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1065 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: circular  
; MOLECULE TYPE: DNA (genomic)  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 1..1062  
US-08-811-949-60

Query Match 100.0%; Score 18; DB 2; Length 1065;

Best Local Similarity 100.0%; Pred. No. 2.9;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCTGAGGGAACAGTGAC 18  
Db 4 TCTGAGGGAACAGTGAC 21

## RESULT 2

US-08-811-949-44  
; Sequence 44, Application US/08811949  
; Patent No. 5840533  
; GENERAL INFORMATION:  
; APPLICANT: NIWA, MINEO  
; APPLICANT: SAITO, YOSHIMASA  
; APPLICANT: SASAKI, HITOSHI  
; APPLICANT: HAYASHI, MASAKO  
; APPLICANT: NOTANI, JOUJI  
; APPLICANT: KOBAYASHI, MASAKAZU  
; TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR  
; NUMBER OF SEQUENCES: 67  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
; ADDRESS: P.C.  
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400  
; CITY: ARLINGTON  
; STATE: VA  
; COUNTRY: USA  
; ZIP: 22202

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/811.949  
FILING DATE: 05-MAR-1997  
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:  
NAME: OBLON, NORMAN F.  
REGISTRATION NUMBER: 24,618  
REFERENCE/DOCKET NUMBER: 18-966-0  
TELEPHONE: 703-413-3000  
TELEFAX: 703-413-2220  
INFORMATION FOR SEQ ID NO: 44:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1068 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: circular  
MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..1065  
US-08-811-949-44

Query Match 100.0%; Score 18; DB 2; Length 1068;  
Best Local Similarity 100.0%; Pred. No. 2.9;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCTGAGGGAACAGTGAC 18  
Db 4 TCTGAGGGAACAGTGAC 21

## RESULT 3

US-08-811-949-46  
; Sequence 46, Application US/08811949  
; Patent No. 5840533  
; GENERAL INFORMATION:  
; APPLICANT: NIWA, MINEO  
; APPLICANT: SAITO, YOSHIMASA

APPLICANT: SASAKI, HITOSHI  
APPLICANT: HAYASHI, MASAKO  
APPLICANT: NOTANI, JOUJI  
APPLICANT: KOBAYASHI, MASAKAZU  
TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR  
NUMBER OF SEQUENCES: 67  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
ADDRESS: P.C.  
STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400  
CITY: ARLINGTON  
STATE: VA  
COUNTRY: USA  
ZIP: 22202

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/811.949  
FILING DATE: 05-MAR-1997  
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:  
NAME: OBLON, NORMAN F.  
REGISTRATION NUMBER: 24,618  
REFERENCE/DOCKET NUMBER: 18-966-0  
TELEPHONE: 703-413-3000  
TELEFAX: 703-413-2220  
INFORMATION FOR SEQ ID NO: 46:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1068 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: circular  
MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..1065  
US-08-811-949-46

Query Match 100.0%; Score 18; DB 2; Length 1068;  
Best Local Similarity 100.0%; Pred. No. 2.9;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCTGAGGGAACAGTGAC 18  
Db 4 TCTGAGGGAACAGTGAC 21

## RESULT 4

US-08-811-949-52  
; Sequence 52, Application US/08811949  
; Patent No. 5840533

GENERAL INFORMATION:

APPLICANT: NIWA, MINEO  
APPLICANT: SAITO, YOSHIMASA  
APPLICANT: SASAKI, HITOSHI  
APPLICANT: HAYASHI, MASAKO  
APPLICANT: NOTANI, JOUJI  
APPLICANT: KOBAYASHI, MASAKAZU  
TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR  
NUMBER OF SEQUENCES: 67  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
ADDRESS: P.C.  
STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400  
CITY: ARLINGTON  
STATE: VA  
COUNTRY: USA  
ZIP: 22202  
COMPUTER READABLE FORM:

;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.30  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/811,949  
;; FILING DATE: 05-MAR-1997  
;; CLASSIFICATION: 435  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: OBLON, NORMAN F.  
;; REGISTRATION NUMBER: 24,618  
;; REFERENCE/DOCKET NUMBER: 18-966-0  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 703-413-3000  
;; TELEFAX: 703-413-2220  
;; INFORMATION FOR SEQ ID NO: 52:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 1068 base pairs  
;; TYPE: nucleic acid  
;; STRANDEDNESS: double  
;; TOPOLOGY: circular  
;; MOLECULE TYPE: DNA (genomic)  
;; FEATURE:  
;; NAME/KEY: CDS  
;; LOCATION: 1..1065  
;; US-08-811-949-52

Query Match 100.0%; Score 18; DB 2; Length 1068;  
Best Local Similarity 100.0%; Pred. No. 2.9;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCTGAGGGAACAGTGAC 18  
|||||  
Db 4 TCTGAGGGAACAGTGAC 21

RESULT 5  
US-08-811-949-58  
;; Sequence 58, Application US/08811949  
;; Patent No. 5840533  
;; GENERAL INFORMATION:  
;; APPLICANT: NIWA, MINEO  
;; APPLICANT: SAITO, YOSHIMASA  
;; APPLICANT: SASAKI, HITOSHI  
;; APPLICANT: HAYASHI, MASAKO  
;; APPLICANT: NOTANI, JOUJI  
;; APPLICANT: KOBAYASHI, MASAKAZU  
;; TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR  
;; NUMBER OF SEQUENCES: 67  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
;; ADDRESSEE: P.C.  
;; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400  
;; CITY: ARLINGTON  
;; STATE: VA  
;; COUNTRY: USA  
;; ZIP: 22202  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.30  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/811,949  
;; FILING DATE: 05-MAR-1997  
;; CLASSIFICATION: 435  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: OBLON, NORMAN F.  
;; REGISTRATION NUMBER: 24,618  
;; REFERENCE/DOCKET NUMBER: 18-966-0  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 703-413-3000  
;; TELEFAX: 703-413-2220

;; INFORMATION FOR SEQ ID NO: 58:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 1068 base pairs  
;; TYPE: nucleic acid  
;; STRANDEDNESS: double  
;; TOPOLOGY: circular  
;; MOLECULE TYPE: DNA (genomic)  
;; FEATURE:  
;; NAME/KEY: CDS  
;; LOCATION: 1..1065  
;; US-08-811-949-58

Query Match 100.0%; Score 18; DB 2; Length 1068;  
Best Local Similarity 100.0%; Pred. No. 2.9;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCTGAGGGAACAGTGAC 18  
|||||  
Db 4 TCTGAGGGAACAGTGAC 21

RESULT 6  
US-08-811-949-64  
;; Sequence 64, Application US/08811949  
;; Patent No. 5840533  
;; GENERAL INFORMATION:  
;; APPLICANT: NIWA, MINEO  
;; APPLICANT: SAITO, YOSHIMASA  
;; APPLICANT: SASAKI, HITOSHI  
;; APPLICANT: HAYASHI, MASAKO  
;; APPLICANT: NOTANI, JOUJI  
;; APPLICANT: KOBAYASHI, MASAKAZU  
;; TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR  
;; NUMBER OF SEQUENCES: 67  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
;; ADDRESSEE: P.C.  
;; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400  
;; CITY: ARLINGTON  
;; STATE: VA  
;; COUNTRY: USA  
;; ZIP: 22202  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.30  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/811,949  
;; FILING DATE: 05-MAR-1997  
;; CLASSIFICATION: 435  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: OBLON, NORMAN F.  
;; REGISTRATION NUMBER: 24,618  
;; REFERENCE/DOCKET NUMBER: 18-966-0  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 703-413-3000  
;; TELEFAX: 703-413-2220  
;; INFORMATION FOR SEQ ID NO: 64:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 1170 base pairs  
;; TYPE: nucleic acid  
;; STRANDEDNESS: double  
;; TOPOLOGY: circular  
;; MOLECULE TYPE: DNA (genomic)  
;; FEATURE:  
;; NAME/KEY: CDS  
;; LOCATION: 1..1167  
;; US-08-811-949-64

Query Match 100.0%; Score 18; DB 2; Length 1170;  
Best Local Similarity 100.0%; Pred. No. 2.9;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCTGAGGAAACAGTGAC 18  
 |||||  
 Db 106 TCTGAGGAAACAGTGAC 123

## RESULT 7

US-08-811-949-66  
 ; Sequence 66, Application US/08811949  
 ; Patent No. 5840533  
 ; GENERAL INFORMATION:  
 ; APPLICANT: NIWA, MINEO  
 ; APPLICANT: SAITO, YOSHIMASA  
 ; APPLICANT: SASAKI, HITOSHI  
 ; APPLICANT: HAYASHI, MASAKO  
 ; APPLICANT: NOTANI, JOUJI  
 ; APPLICANT: KOBAYASHI, MASAKAZU  
 ; TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR  
 ; NUMBER OF SEQUENCES: 67  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
 ; ADDRESSEE: P.C.  
 ; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400  
 ; CITY: ARLINGTON  
 ; STATE: VA  
 ; COUNTRY: USA  
 ; ZIP: 22202

COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/811,949  
 ; FILING DATE: 05-MAR-1997  
 ; CLASSIFICATION: 435  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: OBLON, NORMAN F.  
 ; REGISTRATION NUMBER: 24,618  
 ; REFERENCE/DOCKET NUMBER: 18-966-0  
 ; TELEPHONE: 703-413-3000  
 ; TELEFAX: 703-413-2220  
 ; INFORMATION FOR SEQ ID NO: 66:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 1170 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: double  
 ; TOPOLOGY: circular  
 ; MOLECULE TYPE: DNA (genomic)  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: 1..1167

US-08-811-949-66

Query Match 100.0%; Score 18; DB 2; Length 1170;  
 Best Local Similarity 100.0%; Pred. No. 2.9;  
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCTGAGGAAACAGTGAC 18  
 |||||  
 Db 106 TCTGAGGAAACAGTGAC 123

## RESULT 8

US-08-811-949-48  
 ; Sequence 48, Application US/08811949  
 ; Patent No. 5840533  
 ; GENERAL INFORMATION:  
 ; APPLICANT: NIWA, MINEO  
 ; APPLICANT: SAITO, YOSHIMASA  
 ; APPLICANT: SASAKI, HITOSHI  
 ; APPLICANT: HAYASHI, MASAKO

; APPLICANT: NOTANI, JOUJI  
 ; APPLICANT: KOBAYASHI, MASAKAZU  
 ; TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR  
 ; NUMBER OF SEQUENCES: 67  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
 ; ADDRESSEE: P.C.  
 ; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400  
 ; CITY: ARLINGTON  
 ; STATE: VA  
 ; COUNTRY: USA  
 ; ZIP: 22202  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/811,949  
 ; FILING DATE: 05-MAR-1997  
 ; CLASSIFICATION: 435  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: OBLON, NORMAN F.  
 ; REGISTRATION NUMBER: 24,618  
 ; REFERENCE/DOCKET NUMBER: 18-966-0  
 ; TELEPHONE: 703-413-3000  
 ; TELEFAX: 703-413-2220  
 ; INFORMATION FOR SEQ ID NO: 48:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 1314 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: double  
 ; TOPOLOGY: circular  
 ; MOLECULE TYPE: DNA (genomic)  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: 1..1311  
 ; US-08-811-949-48

Query Match 100.0%; Score 18; DB 2; Length 1314;  
 Best Local Similarity 100.0%; Pred. No. 2.9;  
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCTGAGGAAACAGTGAC 18  
 |||||  
 Db 250 TCTGAGGAAACAGTGAC 267

## RESULT 9

US-08-811-949-50  
 ; Sequence 50, Application US/08811949  
 ; Patent No. 5840533  
 ; GENERAL INFORMATION:  
 ; APPLICANT: NIWA, MINEO  
 ; APPLICANT: SAITO, YOSHIMASA  
 ; APPLICANT: SASAKI, HITOSHI  
 ; APPLICANT: HAYASHI, MASAKO  
 ; APPLICANT: NOTANI, JOUJI  
 ; APPLICANT: KOBAYASHI, MASAKAZU  
 ; TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR  
 ; NUMBER OF SEQUENCES: 67  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
 ; ADDRESSEE: P.C.  
 ; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400  
 ; CITY: ARLINGTON  
 ; STATE: VA  
 ; COUNTRY: USA  
 ; ZIP: 22202

COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible



;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: Patent In Release #1.0, Version #1.30  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/811,949  
;; FILING DATE: 05-MAR-1997  
;; CLASSIFICATION: 435  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: OBLON, NORMAN F.  
;; REGISTRATION NUMBER: 24,618  
;; REFERENCE/DOCKET NUMBER: 18-966-0  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 703-413-3000  
;; TELEFAX: 703-413-2220  
;; INFORMATION FOR SEQ ID NO: 50:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 1314 base pairs  
;; TYPE: nucleic acid  
;; STRANDEDNESS: double  
;; TOPOLOGY: circular  
;; MOLECULE TYPE: DNA (genomic)  
;; FEATURE:  
;; NAME/KEY: CDS  
;; LOCATION: 1..1311  
US-08-811-949-50

Query Match 100.0%; Score 18; DB 2; Length 1314;  
Best Local Similarity 100.0%; Pred. No. 2.9;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCTGAGGGAACAGTGAC 18  
|||||  
Db 250 TCTGAGGGAACAGTGAC 267

RESULT 10  
;; Sequence 54, Application US/08811949  
;; Patent No. 5840533  
;; GENERAL INFORMATION:  
;; APPLICANT: NIWA, MINEO  
;; APPLICANT: SAITO, YOSHIMASA  
;; APPLICANT: SASAKI, HITOSHI  
;; APPLICANT: HAYASHI, MASAKO  
;; APPLICANT: NOTANI, JOUJI  
;; APPLICANT: KOBAYASHI, MASAKAZU  
;; TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR  
;; NUMBER OF SEQUENCES: 67  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
;; ADDRESSEE: P.C.  
;; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400  
;; CITY: ARLINGTON  
;; STATE: VA  
;; COUNTRY: USA  
;; ZIP: 22202  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: Patent In Release #1.0, Version #1.30  
;; CURRENT APPLICATION DATA:  
;; FILING DATE: 05-MAR-1997  
;; CLASSIFICATION: 435  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: OBLON, NORMAN F.  
;; REGISTRATION NUMBER: 24,618  
;; REFERENCE/DOCKET NUMBER: 18-966-0  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 703-413-3000  
;; TELEFAX: 703-413-2220  
;; INFORMATION FOR SEQ ID NO: 54:  
;; SEQUENCE CHARACTERISTICS:

;; LENGTH: 1314 base pairs  
;; TYPE: nucleic acid  
;; STRANDEDNESS: double  
;; TOPOLOGY: circular  
;; MOLECULE TYPE: DNA (genomic)  
;; FEATURE:  
;; NAME/KEY: CDS  
;; LOCATION: 1..1311  
US-08-811-949-54

Query Match 100.0%; Score 18; DB 2; Length 1314;  
Best Local Similarity 100.0%; Pred. No. 2.9;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCTGAGGGAACAGTGAC 18  
|||||  
Db 250 TCTGAGGGAACAGTGAC 267

RESULT 11  
;; Sequence 56, Application US/08811949  
;; Patent No. 5840533  
;; GENERAL INFORMATION:  
;; APPLICANT: NIWA, MINEO  
;; APPLICANT: SAITO, YOSHIMASA  
;; APPLICANT: SASAKI, HITOSHI  
;; APPLICANT: HAYASHI, MASAKO  
;; APPLICANT: NOTANI, JOUJI  
;; APPLICANT: KOBAYASHI, MASAKAZU  
;; TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR  
;; NUMBER OF SEQUENCES: 67  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
;; ADDRESSEE: P.C.  
;; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400  
;; CITY: ARLINGTON  
;; STATE: VA  
;; COUNTRY: USA  
;; ZIP: 22202  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: Patent In Release #1.0, Version #1.30  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/811,949  
;; FILING DATE: 05-MAR-1997  
;; CLASSIFICATION: 435  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: OBLON, NORMAN F.  
;; REGISTRATION NUMBER: 24,618  
;; REFERENCE/DOCKET NUMBER: 18-966-0  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 703-413-3000  
;; TELEFAX: 703-413-2220  
;; INFORMATION FOR SEQ ID NO: 56:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 1314 base pairs  
;; TYPE: nucleic acid  
;; STRANDEDNESS: double  
;; TOPOLOGY: circular  
;; MOLECULE TYPE: DNA (genomic)  
;; FEATURE:  
;; NAME/KEY: CDS  
;; LOCATION: 1..1311  
US-08-811-949-56

Query Match 100.0%; Score 18; DB 2; Length 1314;  
Best Local Similarity 100.0%; Pred. No. 2.9;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCTGAGGGAACAGTGAC 18

```
Db      250 TCTGAGGGAACAGTGAC 267
|||||
RESULT 12
US-08-811-949-62
; Sequence 62, Application US/08811949
; Patent No. 5840533
; GENERAL INFORMATION:
; APPLICANT: NIWA, MINEO
; APPLICANT: SAITO, YOSHIMASA
; APPLICANT: SASAKI, HITOSHI
; APPLICANT: HAYASHI, MASAKO
; APPLICANT: NOTANI, JOUJI
; APPLICANT: KOBAYASHI, MASAKAZU
; TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR
; NUMBER OF SEQUENCES: 67
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; ADDRESSEE: P.C.
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/811,949
; FILING DATE: 05-MAR-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 18-966-0
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 62:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1419 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1416
US-08-811-949-62

Query Match      100.0%; Score 18; DB 2; Length 1419;
Best Local Similarity 100.0%; Pred. No. 3;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 TCTGAGGGAACAGTGAC 18
|||||
Db      355 TCTGAGGGAACAGTGAC 372
|||||

RESULT 13
5200340-5
; Patent No. 5200340
; APPLICANT: FOSTER, DONALD C.; MULVIHILL, EILEEN R.; O'HARA,
; PATRICK J.; PINGEL, KURT; YOSHITAKE, SHINJI
; TITLE OF INVENTION: THROMBIN-ACTIVATED TISSUE PLASMINOGEN
; ACTIVATORS
; NUMBER OF SEQUENCES: 34
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/53,412
; FILING DATE: 22-MAY-1987
```

```
;SEQ ID NO:5:
; LENGTH: 1724
5200340-5

Query Match      100.0%; Score 18; DB 6; Length 1724;
Best Local Similarity 100.0%; Pred. No. 3;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 TCTGAGGGAACAGTGAC 18
|||||
Db      640 TCTGAGGGAACAGTGAC 657
|||||

RESULT 14
5200340-1
; Patent No. 5200340
; APPLICANT: FOSTER, DONALD C.; MULVIHILL, EILEEN R.; O'HARA,
; PATRICK J.; PINGEL, KURT; YOSHITAKE, SHINJI
; TITLE OF INVENTION: THROMBIN-ACTIVATED TISSUE PLASMINOGEN
; ACTIVATORS
; NUMBER OF SEQUENCES: 34
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/53,412
; FILING DATE: 22-MAY-1987
;SEQ ID NO:1:
; LENGTH: 1738
5200340-1

Query Match      100.0%; Score 18; DB 6; Length 1738;
Best Local Similarity 100.0%; Pred. No. 3;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 TCTGAGGGAACAGTGAC 18
|||||
Db      640 TCTGAGGGAACAGTGAC 657
|||||

RESULT 15
US-08-814-412-10
; Sequence 10, Application US/08814412
; Patent No. 6150141
; GENERAL INFORMATION:
; APPLICANT: Jarrell Ph.D., Kevin A.
; TITLE OF INVENTION: Intron-Mediated Techniques and Reagents
; NUMBER OF SEQUENCES: 46
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Choate, Hall & Stewart
; STREET: 53 State Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/814,412
; FILING DATE: 11-MAR-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Jarrell Ph.D., Brenda H.
; REGISTRATION NUMBER: 39,223
; REFERENCE/DOCKET NUMBER: 0079571-0040
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617 248 5000
; TELEFAX: 617 248 4000
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1848 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
```

```
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; IMMEDIATE SOURCE:
; CLONE: amplified t-PA clone
US-08-814-412-10

Query Match      100.0%; Score 18; DB 3; Length 1848;
Best Local Similarity 100.0%; Pred. No. 3;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 TCTGAGGGGAACAGTGAC 18
        |||||
Db      696 TCTGAGGGGAACAGTGAC 713

Search completed: August 19, 2003, 08:35:04
Job time : 4.40873 secs
```

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 19, 2003, 04:46:44 ; Search time 5.07937 Seconds  
(without alignments)  
7930.701 Million cell updates/sec

Title: US-09-987-455-1

Perfect score: 18

Sequence: 1 tctgagggaacagtgc 18

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 1504479 seqs, 1118970152 residues

Total number of hits satisfying chosen parameters: 3008958

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA:\*

- 1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/ptodata/2/pubpna/FCT\_NEW\_PUB.seq:\*
- 3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:\*
- 4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:\*
- 5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:\*
- 6: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*
- 7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:\*
- 8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*
- 9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq:\*
- 10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq:\*
- 11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq:\*
- 12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:\*
- 13: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq:\*
- 14: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq:\*
- 15: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:\*
- 16: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:\*
- 17: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	18	100.0	18	11	US-09-987-457-5
2	18	100.0	18	11	US-09-987-455-1
3	18	100.0	42	11	US-09-987-457-6
4	18	100.0	42	11	US-09-987-455-22
5	18	100.0	1065	11	US-09-987-455-4
6	18	100.0	1065	11	US-09-987-455-7
7	18	100.0	1128	11	US-09-987-455-2
8	18	100.0	1128	11	US-09-987-455-5
9	18	100.0	1415	14	US-10-198-846-12748
10	18	100.0	1689	9	US-09-969-271-6
11	18	100.0	2509	14	US-10-193-656-7
12	18	100.0	2519	9	US-09-969-271-5
13	18	100.0	2641	10	US-09-974-298-144
14	15.4	85.6	139	10	US-09-878-574-237
15	15.4	85.6	292	9	US-09-864-761-18405
16	15.4	85.6	352	10	US-09-960-352-6540

17	15.4	85.6	403	10	US-09-983-965-2988	Sequence 2988, Ap
18	15.4	85.6	403	10	US-09-983-965-4942	Sequence 4942, Ap
c 19	15.4	85.6	466	9	US-09-864-761-1647	Sequence 1647, Ap
c 20	15.4	85.6	530	13	US-10-027-632-39133	Sequence 39133, A
c 21	15.4	85.6	675	13	US-10-027-632-243054	Sequence 243054, A
c 22	15.4	85.6	767	13	US-10-027-632-157405	Sequence 157405, A
c 23	15.4	85.6	818	13	US-10-027-632-166380	Sequence 166380, A
c 24	15.4	85.6	936	13	US-10-027-632-166215	Sequence 166215, A
c 25	15.4	85.6	936	13	US-10-027-632-166216	Sequence 166216, A
c 26	15.4	85.6	1184	13	US-10-027-632-101376	Sequence 101376, A
c 27	15.4	85.6	1225	13	US-10-027-632-85099	Sequence 85099, A
c 28	15.4	85.6	1360	12	US-10-017-161-947	Sequence 947, App
c 29	15.4	85.6	1459	9	US-09-759-143-174	Sequence 174, App
c 30	15.4	85.6	1459	9	US-09-780-669-174	Sequence 174, App
c 31	15.4	85.6	1459	9	US-09-030-606-174	Sequence 174, App
c 32	15.4	85.6	1459	9	US-09-822-827-174	Sequence 174, App
c 33	15.4	85.6	1459	9	US-09-115-453-174	Sequence 174, App
c 34	15.4	85.6	1459	10	US-09-232-880-174	Sequence 174, App
c 35	15.4	85.6	1459	10	US-09-895-793-174	Sequence 174, App
c 36	15.4	85.6	1459	10	US-09-895-814-174	Sequence 174, App
c 37	15.4	85.6	1459	13	US-10-012-896-174	Sequence 174, App
c 38	15.4	85.6	1459	14	US-10-010-940-174	Sequence 174, App
c 39	15.4	85.6	1590	9	US-09-888-615-3	Sequence 3, Appli
c 40	15.4	85.6	1617	10	US-09-917-800A-1699	Sequence 1699, Ap
c 41	15.4	85.6	1617	12	US-10-205-194-175	Sequence 175, App
c 42	15.4	85.6	1901	13	US-10-027-632-266063	Sequence 266063, A
c 43	15.4	85.6	2392	12	US-09-813-432-23	Sequence 23, Appl
c 44	15.4	85.6	3772	9	US-09-989-722-61	Sequence 61, Appl
c 45	15.4	85.6	3772	9	US-09-989-723-61	Sequence 61, Appl

#### ALIGNMENTS

#### RESULT 1

US-09-987-457-5  
; Sequence 5, Application US/09987457  
; Publication No. US20030013150A1  
; GENERAL INFORMATION:  
; APPLICANT: Manosroi, Aranya  
; APPLICANT: Manosroi, Jiradej  
; APPLICANT: Tayapiwatana, Chatchai  
; APPLICANT: Goetz, Friedrich  
; APPLICANT: Werner, Rolf-Guenther  
; TITLE OF INVENTION: Methods for Large Scale Protein Production in Prokaryotes  
; FILE REFERENCE: 0652.2180001  
; CURRENT APPLICATION NUMBER: US/09/987,457  
; CURRENT FILING DATE: 2001-11-14  
; PRIOR APPLICATION NUMBER: 60/268,573  
; PRIOR FILING DATE: 2001-02-15  
; PRIOR APPLICATION NUMBER: GB 00 27 782.2  
; PRIOR FILING DATE: 2000-11-14  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 5  
; LENGTH: 18  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: coding  
; OTHER INFORMATION: sequence of the N-terminal part of K2S molecule  
US-09-987-457-5

Query Match 100.0%; Score 18; DB 11; Length 18;

Best Local Similarity 100.0%; Pred. No. 7.1; Mismatches 0; Indels 0; Gaps 0;

Matches 18; Conservative 0;

Qy 1 TCTGAGGGAACAGTGCAC 18

Db 1 TCTGAGGGAACAGTGCAC 18

#### RESULT 2

US-09-987-455-1  
; Sequence 1, Application US/09987455  
; Publication No. US20030049729A1  
; GENERAL INFORMATION:  
; APPLICANT: Aranya Manosroi  
; APPLICANT: Jiradej Manosroi  
; APPLICANT: Chatchai Tayapiwatana  
; APPLICANT: Friedrich Goetz  
; APPLICANT: Rolf-Guenther Werner  
; TITLE OF INVENTION: Methods for Large Scale Production of Recombinant  
; FILE REFERENCE: 0652.2190001  
; CURRENT APPLICATION NUMBER: US/09/987,455  
; PRIOR FILING DATE: 2001-11-14  
; PRIOR APPLICATION NUMBER: 60/268,574  
; PRIOR FILING DATE: 2001-02-15  
; PRIOR APPLICATION NUMBER: GB 0027779.8  
; PRIOR FILING DATE: 2000-11-14  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: coding  
; OTHER INFORMATION: sequence of N-terminal part of K2S protein  
US-09-987-455-1  
Query Match 100.0%; Score 18; DB 11; Length 18;  
Best Local Similarity 100.0%; Pred. No. 7.1; Mismatches 0; Indels 0; Gaps 0;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 TCTGAGGAAACAGTGAC 18  
|||||  
DB 1 TCTGAGGAAACAGTGAC 18  
|||||  
RESULT 3  
US-09-987-457-6  
; Sequence 6, Application US/09987457  
; Publication No. US20030013150A1  
; GENERAL INFORMATION:  
; APPLICANT: Manosroi, Aranya  
; APPLICANT: Manosroi, Jiradej  
; APPLICANT: Tayapiwatana, Chatchai  
; APPLICANT: Goetz, Friedrich  
; APPLICANT: Werner, Rolf-Guenther  
; TITLE OF INVENTION: Methods for Large Scale Protein Production in Prokaryotes  
; FILE REFERENCE: 0652.2180001  
; CURRENT APPLICATION NUMBER: US/09/987,457  
; CURRENT FILING DATE: 2001-11-14  
; PRIOR APPLICATION NUMBER: 60/268,573  
; PRIOR FILING DATE: 2001-02-15  
; PRIOR APPLICATION NUMBER: GB 00 27 782.2  
; PRIOR FILING DATE: 2000-11-14  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 6  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: oligonucleotide  
; OTHER INFORMATION: sequence  
US-09-987-457-6  
Query Match 100.0%; Score 18; DB 11; Length 42;  
Best Local Similarity 100.0%; Pred. No. 7.2; Mismatches 0; Indels 0; Gaps 0;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 TCTGAGGAAACAGTGAC 18  
|||||

Db 25 TCTGAGGAAACAGTGAC 42  
|||||  
RESULT 4  
US-09-987-455-22  
; Sequence 22, Application US/09987455  
; Publication No. US20030049729A1  
; GENERAL INFORMATION:  
; APPLICANT: Aranya Manosroi  
; APPLICANT: Jiradej Manosroi  
; APPLICANT: Chatchai Tayapiwatana  
; APPLICANT: Friedrich Goetz  
; APPLICANT: Rolf-Guenther Werner  
; TITLE OF INVENTION: Methods for Large Scale Production of Recombinant  
; FILE REFERENCE: 0652.2190001  
; CURRENT APPLICATION NUMBER: US/09/987,455  
; CURRENT FILING DATE: 2001-11-14  
; PRIOR APPLICATION NUMBER: 60/268,574  
; PRIOR FILING DATE: 2001-02-15  
; PRIOR APPLICATION NUMBER: GB 0027779.8  
; PRIOR FILING DATE: 2000-11-14  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 22  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: primer  
US-09-987-455-22  
Query Match 100.0%; Score 18; DB 11; Length 42;  
Best Local Similarity 100.0%; Pred. No. 7.2; Mismatches 0; Indels 0; Gaps 0;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 TCTGAGGAAACAGTGAC 18  
|||||  
Db 25 TCTGAGGAAACAGTGAC 42  
|||||  
RESULT 5  
US-09-987-455-4  
; Sequence 4, Application US/09987455  
; Publication No. US20030049729A1  
; GENERAL INFORMATION:  
; APPLICANT: Aranya Manosroi  
; APPLICANT: Jiradej Manosroi  
; APPLICANT: Chatchai Tayapiwatana  
; APPLICANT: Friedrich Goetz  
; APPLICANT: Rolf-Guenther Werner  
; TITLE OF INVENTION: Methods for Large Scale Production of Recombinant  
; FILE REFERENCE: 0652.2190001  
; CURRENT APPLICATION NUMBER: US/09/987,455  
; CURRENT FILING DATE: 2001-11-14  
; PRIOR APPLICATION NUMBER: 60/268,574  
; PRIOR FILING DATE: 2001-02-15  
; PRIOR APPLICATION NUMBER: GB 0027779.8  
; PRIOR FILING DATE: 2000-11-14  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 4  
; LENGTH: 1065  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: coding  
; OTHER INFORMATION: sequence for K2S protein  
US-09-987-455-4  
Query Match 100.0%; Score 18; DB 11; Length 1065;  
Best Local Similarity 100.0%; Pred. No. 7.6;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCTGAGGGAACAGTGAC 18  
| | | | | | | | | | | | | | | | | |  
Db 1 TCTGAGGGAACAGTGAC 18

## RESULT 6

US-09-987-455-7  
; Sequence 7, Application US/09987455  
; Publication No. US20030049729A1  
; GENERAL INFORMATION:  
; APPLICANT: Aranya Manosroi  
; APPLICANT: Jiradej Manosroi  
; APPLICANT: Chatchai Tayapiwatana  
; APPLICANT: Friedrich Goetz  
; APPLICANT: Rolf-Guenther Werner  
; TITLE OF INVENTION: Methods for Large Scale Production of Recombinant  
; TITLE OF INVENTION: DNA-Derived tPA or K2S Molecules  
; FILE REFERENCE: 0652.2190001  
; CURRENT APPLICATION NUMBER: US/09/987,455  
; CURRENT FILING DATE: 2001-11-14  
; PRIOR APPLICATION NUMBER: 60/268,574  
; PRIOR FILING DATE: 2001-02-15  
; PRIOR APPLICATION NUMBER: GB 0027779.8  
; PRIOR FILING DATE: 2000-11-14  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 7  
; LENGTH: 1065  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: coding  
; OTHER INFORMATION: sequence for K2S protein  
US-09-987-455-7

Query Match 100.0%; Score 18; DB 11; Length 1065;  
Best Local Similarity 100.0%; Pred. No. 7.6;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCTGAGGGAACAGTGAC 18  
| | | | | | | | | | | | | | | | | |  
Db 1 TCTGAGGGAACAGTGAC 18

## RESULT 7

US-09-987-455-2  
; Sequence 2, Application US/09987455  
; Publication No. US20030049729A1  
; GENERAL INFORMATION:  
; APPLICANT: Aranya Manosroi  
; APPLICANT: Jiradej Manosroi  
; APPLICANT: Chatchai Tayapiwatana  
; APPLICANT: Friedrich Goetz  
; APPLICANT: Rolf-Guenther Werner  
; TITLE OF INVENTION: Methods for Large Scale Production of Recombinant  
; TITLE OF INVENTION: DNA-Derived tPA or K2S Molecules  
; FILE REFERENCE: 0652.2190001  
; CURRENT APPLICATION NUMBER: US/09/987,455  
; CURRENT FILING DATE: 2001-11-14  
; PRIOR APPLICATION NUMBER: 60/268,574  
; PRIOR FILING DATE: 2001-02-15  
; PRIOR APPLICATION NUMBER: GB 0027779.8  
; PRIOR FILING DATE: 2000-11-14  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 1128  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: coding  
US-09-987-455-2

; OTHER INFORMATION: sequence for OmpA-K2S fusion protein  
US-09-987-455-2

Query Match 100.0%; Score 18; DB 11; Length 1128;  
Best Local Similarity 100.0%; Pred. No. 7.6;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCTGAGGGAACAGTGAC 18  
| | | | | | | | | | | | | | | | | |  
Db 67 TCTGAGGGAACAGTGAC 84

## RESULT 8

US-09-987-455-5  
; Sequence 5, Application US/09987455  
; Publication No. US20030049729A1  
; GENERAL INFORMATION:  
; APPLICANT: Aranya Manosroi  
; APPLICANT: Jiradej Manosroi  
; APPLICANT: Chatchai Tayapiwatana  
; APPLICANT: Friedrich Goetz  
; APPLICANT: Rolf-Guenther Werner  
; TITLE OF INVENTION: Methods for Large Scale Production of Recombinant  
; TITLE OF INVENTION: DNA-Derived tPA or K2S Molecules  
; FILE REFERENCE: 0652.2190001  
; CURRENT APPLICATION NUMBER: US/09/987,455  
; CURRENT FILING DATE: 2001-11-14  
; PRIOR APPLICATION NUMBER: 60/268,574  
; PRIOR FILING DATE: 2001-02-15  
; PRIOR APPLICATION NUMBER: GB 0027779.8  
; PRIOR FILING DATE: 2000-11-14  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 5  
; LENGTH: 1128  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: coding  
; OTHER INFORMATION: sequence for OmpA-K2S fusion protein  
US-09-987-455-5

Query Match 100.0%; Score 18; DB 11; Length 1128;  
Best Local Similarity 100.0%; Pred. No. 7.6;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCTGAGGGAACAGTGAC 18  
| | | | | | | | | | | | | | | | | |  
Db 67 TCTGAGGGAACAGTGAC 84

## RESULT 9

US-10-198-846-12748  
; Sequence 12748, Application US/10198846  
; Publication No. US2003009974A1  
; GENERAL INFORMATION:  
; APPLICANT: Lillie, James  
; APPLICANT: Xu, Yongyao  
; APPLICANT: Wang, Youzhen  
; APPLICANT: Steinmann, Kathleen  
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS  
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND  
; TITLE OF INVENTION: THERAPY OF BREAST CANCER  
; FILE REFERENCE: MFI-049  
; CURRENT APPLICATION NUMBER: US/10/198,846  
; CURRENT FILING DATE: 2002-07-18  
; PRIOR APPLICATION NUMBER: 60/306,220  
; PRIOR FILING DATE: 2001-07-18  
; NUMBER OF SEQ ID NOS: 14084  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 12748  
; LENGTH: 1415  
; TYPE: DNA  
US-10-198-846-12748

```
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1, 2, 3, 1414, 1415
; OTHER INFORMATION: n = A,T,C or G
US-10-198-846-12748

Query Match      100.0%; Score 18; DB 14; Length 1415;
Best Local Similarity 100.0%; Pred. No. 7.7;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCTGAGGGAACAGTGAC 18
    |||||
Db 906 TCTGAGGGAACAGTGAC 923
    |||||

RESULT 10
US-09-969-271-6
; Sequence 6, Application US/09969271
; Patent No. US20020098179A1
; GENERAL INFORMATION:
; APPLICANT: Pfizer Inc. (All designated States except GB and EP (GB));
; APPLICANT: Pfizer Limited (GB and EP (GB) only)
; TITLE OF INVENTION: Pharmaceutical Combinations
; FILE REFERENCE: PCS10951APME
; CURRENT APPLICATION NUMBER: US/09/969,271
; CURRENT FILING DATE: 2001-10-01
; PRIOR APPLICATION NUMBER: GB 0025473.0
; PRIOR FILING DATE: 2000-10-17
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 1689
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-969-271-6

Query Match      100.0%; Score 18; DB 9; Length 1689;
Best Local Similarity 100.0%; Pred. No. 7.7;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCTGAGGGAACAGTGAC 18
    |||||
Db 625 TCTGAGGGAACAGTGAC 642
    |||||

RESULT 11
US-10-193-656-7
; Sequence 7, Application US/10193656
; Publication No. US20030096733A1
; GENERAL INFORMATION:
; APPLICANT: NY, Tor
; APPLICANT: HOLMDEHL, Rikard
; APPLICANT: Li, Jinan
; TITLE OF INVENTION: NOVEL DRUG TARGETS FOR ARTHRITIS
; FILE REFERENCE: 3810/J577-US3
; CURRENT APPLICATION NUMBER: US/10/193,656
; CURRENT FILING DATE: 2002-07-10
; PRIOR APPLICATION NUMBER: US 60/304,461
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US 60/304,490
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US 60/305,182
; PRIOR FILING DATE: 2001-07-13
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
; LENGTH: 2509
; TYPE: DNA
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: GenBank / X07393
; DATABASE ENTRY DATE: 1995-03-27

; RELEVANT RESIDUES: (1)..(2509)
US-10-193-656-7

Query Match      100.0%; Score 18; DB 14; Length 2509;
Best Local Similarity 100.0%; Pred. No. 7.7;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCTGAGGGAACAGTGAC 18
    |||||
Db 678 TCTGAGGGAACAGTGAC 695
    |||||

RESULT 12
US-09-969-271-5
; Sequence 5, Application US/09969271
; Patent No. US20020098179A1
; GENERAL INFORMATION:
; APPLICANT: Pfizer Inc. (All designated States except GB and EP (GB));
; APPLICANT: Pfizer Limited (GB and EP (GB) only)
; TITLE OF INVENTION: Pharmaceutical Combinations
; FILE REFERENCE: PCS10951APME
; CURRENT APPLICATION NUMBER: US/09/969,271
; CURRENT FILING DATE: 2001-10-01
; PRIOR APPLICATION NUMBER: GB 0025473.0
; PRIOR FILING DATE: 2000-10-17
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 2519
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-969-271-5

Query Match      100.0%; Score 18; DB 9; Length 2519;
Best Local Similarity 100.0%; Pred. No. 7.7;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCTGAGGGAACAGTGAC 18
    |||||
Db 701 TCTGAGGGAACAGTGAC 718
    |||||

RESULT 13
US-09-974-298-144
; Sequence 144, Application US/09974298
; Patent No. US20020156263A1
; GENERAL INFORMATION:
; APPLICANT: Chen, Hwei-Mei
; TITLE OF INVENTION: GENES EXPRESSED IN BREAST CANCER
; FILE REFERENCE: PA-0037 P
; CURRENT APPLICATION NUMBER: US/09/974,298
; CURRENT FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: 60/238,331
; PRIOR FILING DATE: 2000-05-10
; NUMBER OF SEQ ID NOS: 194
; SOFTWARE: PERL Program
; SEQ ID NO 144
; LENGTH: 2641
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20020156263A1 1001470CB1
; NAME/KEY: unsure
; LOCATION: 2635
; OTHER INFORMATION: a, t, c, g, or other
US-09-974-298-144

Query Match      100.0%; Score 18; DB 10; Length 2641;
Best Local Similarity 100.0%; Pred. No. 7.7;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCTGAGGGAACAGTGAC 18
```



Db 821 TCTGAGGGAACAGTGAC 838  
|||||

## RESULT 14

US-09-878-574-237  
; Sequence 237, Application US/09878574  
; Patent No. US20020110548A1  
; GENERAL INFORMATION:  
; APPLICANT: Byrum, Joseph R.  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Thompson, Michael D.  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with  
; TITLE OF INVENTION: Plants  
; FILE REFERENCE: 38-21(15401)B  
; CURRENT APPLICATION NUMBER: US/09/878,574  
; CURRENT FILING DATE: 2001-12-21  
; PRIOR APPLICATION NUMBER: 09/333,535  
; PRIOR FILING DATE: 1999-06-14  
; NUMBER OF SEQ ID NOS: 15775  
; SEQ ID NO 237  
; LENGTH: 139  
; TYPE: DNA  
; ORGANISM: Glycine max  
; OTHER INFORMATION: Clone ID: LIB3028-053-Q1-B1-B8  
US-09-878-574-237

Query Match 85.6%; Score 15.4; DB 10; Length 139;

Best Local Similarity 94.1%; Pred. No. 1.6e+02;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TCTGAGGGAACAGTGA 17  
|||||

Db 74 TCTGAGGGAACAGTGA 90  
|||||

## RESULT 15

US-09-864-761-18405/c  
; Sequence 18405, Application US/09864761  
; Patent No. US20020048763A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharon G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; APPLICANT: Chen, Wensheng  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY  
; FILE REFERENCE: Aecmica-X-1  
; CURRENT APPLICATION NUMBER: US/09/864,761  
; CURRENT FILING DATE: 2001-05-23  
; PRIOR APPLICATION NUMBER: US 60/180,312  
; PRIOR FILING DATE: 2000-02-04  
; PRIOR APPLICATION NUMBER: US 60/207,456  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: US 09/632,366  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: GB 24263.6  
; PRIOR FILING DATE: 2000-10-04  
; PRIOR APPLICATION NUMBER: US 60/236,359  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: PCT/US01/00666  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00667  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00664  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00669  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00665  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00668  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00663

; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00662  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00661  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00670  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: US 60/234,687  
; PRIOR FILING DATE: 2000-09-21  
; PRIOR APPLICATION NUMBER: US 09/608,408  
; PRIOR FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: US 09/774,203  
; PRIOR FILING DATE: 2001-01-29  
; NUMBER OF SEQ ID NOS: 49117  
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1  
; SEQ ID NO 18405  
; LENGTH: 292  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: MAP TO AP000459.2  
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 20  
; OTHER INFORMATION: EXPRESSED IN HEL100, SIGNAL = 1e+02  
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 60  
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 6.5  
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 25  
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 14  
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 4.2  
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 61  
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 23  
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 50  
; OTHER INFORMATION: EST HUMAN HIT: BE735581.1, EVALUE 7.00e-82  
; OTHER INFORMATION: SWISSPROT HIT: P02552, EVALUE 6.00e-41  
; OTHER INFORMATION: NT HIT: AF081484.1, EVALUE 4.00e-82  
US-09-864-761-18405

Query Match 85.6%; Score 15.4; DB 9; Length 292;  
Best Local Similarity 94.1%; Pred. No. 1.6e+02;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TCTGAGGGAACAGTGA 17  
|||||

Db 77 TCTGAGGGAACAGTGA 61  
|||||

Search completed: August 19, 2003, 14:22:19  
Job time : 6.07937 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 18, 2003, 22:44:33 ; Search time 52.6587 Seconds  
(without alignments)  
8307.845 Million cell updates/sec

Title: US-09-987-455-1

Perfect score: 18

Sequence: 1 tctgagggaacagtgc 18

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_htc:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_htc:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: em\_gss\_hum:\*  
18: em\_gss\_inv:\*  
19: em\_gss\_pln:\*  
20: em\_gss\_vrt:\*  
21: em\_gss\_fun:\*  
22: em\_gss\_mam:\*  
23: em\_gss\_mus:\*  
24: em\_gss\_pro:\*  
25: em\_gss\_rod:\*  
26: em\_gss\_phg:\*  
27: em\_gss\_vrl:\*  
28: gb\_gesl:\*  
29: gb\_ges2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	18	100.0	241	9	AA376531
2	18	100.0	335	10	BE158218
3	18	100.0	374	10	BE173847
4	18	100.0	378	10	BE173842

5	18	100.0	476	10	BE259860
6	18	100.0	519	14	CB149125
7	18	100.0	582	10	BE386276
8	18	100.0	612	10	BE386111
9	18	100.0	639	9	AU117078
10	18	100.0	644	10	BE477932
11	18	100.0	682	10	BE276291
12	18	100.0	690	10	BE895982
13	18	100.0	711	10	BE476511
14	18	100.0	725	12	BM042154
15	18	100.0	761	12	BI771956
16	18	100.0	778	10	BE408133
17	18	100.0	778	12	BE762705
18	18	100.0	798	10	BE281523
19	18	100.0	811	12	BI460599
20	18	100.0	818	9	AU124602
21	18	100.0	831	13	BU156867
22	18	100.0	835	10	BE281371
23	18	100.0	841	13	BU157242
24	18	100.0	843	9	AU135634
25	18	100.0	843	13	BQ878473
26	18	100.0	845	13	BQ877849
27	18	100.0	856	10	BE743550
28	18	100.0	858	10	BE281843
29	18	100.0	858	13	BQ436922
30	18	100.0	859	13	BU157686
31	18	100.0	861	9	AU133204
32	18	100.0	861	13	BQ681189
33	18	100.0	868	12	BE762691
34	18	100.0	872	10	BE730286
35	18	100.0	872	13	BQ676711
36	18	100.0	872	13	BQ879010
37	18	100.0	877	9	AU131845
38	18	100.0	877	13	BQ682391
39	18	100.0	877	13	BU557319
40	18	100.0	879	13	BU179560
41	18	100.0	886	13	BU160825
42	18	100.0	887	13	BU145169
43	18	100.0	888	10	BF026008
44	18	100.0	898	12	BI757143
45	18	100.0	899	10	BE277041

#### ALIGNMENTS

RESULT 1  
AA376531

LOCUS

DEFINITION

EST88963 HSC172 cells

plasmidogen activator,

tissue type (GB:M15518),

mRNA sequence.

AA376531

VERSION

AA376531.1

GI:2028850

EST.

KEYWORDS

SOURCE

Homosapiens

ORGANISM

REFERENCE

AUTHORS

AA376531 241 bp mRNA linear EST 21-APR-1997  
EST88963 HSC172 cells II Homo sapiens cDNA 5' end similar to  
plasmidogen activator, tissue type (GB:M15518), mRNA sequence.

AA376531

VERSION

AA376531.1

EST.

KEYWORDS

SOURCE

Homosapiens

ORGANISM

REFERENCE

AUTHORS

AA376531 241 bp mRNA linear EST 21-APR-1997  
EST88963 HSC172 cells II Homo sapiens cDNA 5' end similar to  
plasmidogen activator, tissue type (GB:M15518), mRNA sequence.

AA376531

VERSION

AA376531.1

EST.

KEYWORDS

SOURCE

Homosapiens

ORGANISM

REFERENCE

AUTHORS

AA376531 241 bp mRNA linear EST 21-APR-1997  
EST88963 HSC172 cells II Homo sapiens cDNA 5' end similar to  
plasmidogen activator, tissue type (GB:M15518), mRNA sequence.

Wei,Y.F., Wing,J., Xu,C., Yu,G.L., Ruben,S.M., Dillion,P.J., Fannon  
,M.R., Rosen,C.A., Haseltine,W.A., Fields,C., Fraser,C.M. and  
Venter,J.C.

Initial assessment of human gene diversity and expression patterns  
based upon 83 million nucleotides of cDNA sequence

Nature 377 (6547 Suppl), 3-174 (1995)

96026280

PUBMED

COMMENT

7566098

Other ESTs: THC167250

Contact: Kerlavage, AR

Bioinformatics

The Institute for Genomic Research

9712 Medical Center Drive, Rockville, MD 20850 USA

Tel: 3018699056

Fax: 3018699423

Email: arkerlav@tigr.org

For clone availability, additional sequence and expression

information related to this EST, please check the TIGR Human Gene

Index (<http://www.tigr.org/cdb/hgi/hgi.html>)

Seq primer: M13 Reverse.

Location/Qualifiers

1..241

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="ATCC (inhost):180988"

/db\_xref="taxon:9606"

/cell\_type="fibroblast"

/cell\_line="HSC172 (60PDL)"

/dev\_stage="fetal"

/clone\_lib="HSC172 cells II"

/note="Organ: lung; Vector: pBluescript SK-; Site\_1: EcoRI

/Site\_2: XhoI"

54 a 70 c 68 g 46 t 3 others

BASE COUNT

ORIGIN

Query Match

Best Local Similarity 100.0%; Score 18; DB 9; Length 241;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCTGAGGGAAACAGTGAC 18

Db 1 TCTGAGGGAAACAGTGAC 18

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

Fax: +55-11-2707001

Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome

Project. This entry can be seen in the following URL

(<http://www.ludwig.org.br/scripts/gethtml2.pl?tl=et2-MR2-HT0380-310>)

100-102-g08&t3=2000-01-31&t4=1)

Seq primer: puc 18 forward

High quality sequence stop: 335.

FEATURES

source

1..335

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/dev\_stage="Adult"

/clone\_lib="HT0380"

/note="Organ: head neck; Vector: puc18; Site\_1: SmaI;

Site\_2: SmaI; A mini-library was made by cloning products

derived from ORESTES PCR (U.S. Letters Patent application

No. 196,716 - Ludwig Institute for Cancer Research)

profiles into the pUC 18 vector. Reverse transcription of

tissue mRNA and cDNA amplification were performed under

low stringency conditions."

51 a 102 c 89 g 93 t

BASE COUNT

ORIGIN

Query Match

Best Local Similarity 100.0%; Score 18; DB 10; Length 335;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCTGAGGGAAACAGTGAC 18

Db 31 TCTGAGGGAAACAGTGAC 14

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

100-102-g08&t3=2000-01-31&t4=1)

Seq primer: puc 18 forward

High quality sequence stop: 335.

Location/Qualifiers

1..335

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/dev\_stage="Adult"

/clone\_lib="HT0380"

/note="Organ: head neck; Vector: puc18; Site\_1: SmaI;

Site\_2: SmaI; A mini-library was made by cloning products

derived from ORESTES PCR (U.S. Letters Patent application

No. 196,716 - Ludwig Institute for Cancer Research)

profiles into the pUC 18 vector. Reverse transcription of

tissue mRNA and cDNA amplification were performed under

low stringency conditions."

51 a 102 c 89 g 93 t

BASE COUNT

ORIGIN

Query Match

Best Local Similarity 100.0%; Score 18; DB 10; Length 335;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCTGAGGGAAACAGTGAC 18

Db 31 TCTGAGGGAAACAGTGAC 14

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

100-102-g08&t3=2000-01-31&t4=1)

Seq primer: puc 18 forward

High quality sequence stop: 335.

Location/Qualifiers

1..335

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/dev\_stage="Adult"

/clone\_lib="HT0380"

/note="Organ: head neck; Vector: puc18; Site\_1: SmaI;

Site\_2: SmaI; A mini-library was made by cloning products

derived from ORESTES PCR (U.S. Letters Patent application

No. 196,716 - Ludwig Institute for Cancer Research)

profiles into the pUC 18 vector. Reverse transcription of

tissue mRNA and cDNA amplification were performed under

low stringency conditions."

51 a 102 c 89 g 93 t

BASE COUNT

ORIGIN

Query Match

Best Local Similarity 100.0%; Score 18; DB 10; Length 335;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCTGAGGGAAACAGTGAC 18

Db 31 TCTGAGGGAAACAGTGAC 14

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="HT0567"
/Note="Organ: head neck; Vector: puc18; Site 1: SmaI;
Site 2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
BASE COUNT      68 a 116 c 123 g 67 t
ORIGIN

Query Match      100.0%; Score 18; DB 10; Length 374;
Best Local Similarity 100.0%; Pred. No. 4e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCTGAGGGAACAGTGAC 18
|||||
Db 157 TCTGAGGGAACAGTGAC 174

RESULT 4
BE173842
LOCUS      CM0-HT0567-010300-256-e04 HT0567 Homo sapiens cDNA, mRNA sequence.
DEFINITION
ACCESSION  BE173842
VERSION    BE173842.1 GI:8636568
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 378)
Dias Neto,B., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deoliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
10737800
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?t1=et2=CM0-HT0567-010
300-256-e04&t3=2000-03-01&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 378.
Location/Qualifiers
1..378
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="HT0567"
/Note="Organ: head neck; Vector: puc18; Site 1: SmaI;
Site 2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of

```

```

tissue mRNA and cDNA amplification were performed under
low stringency conditions."
BASE COUNT      68 a 116 c 126 g 68 t
ORIGIN

Query Match      100.0%; Score 18; DB 10; Length 378;
Best Local Similarity 100.0%; Pred. No. 4.1e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCTGAGGGAACAGTGAC 18
|||||
Db 161 TCTGAGGGAACAGTGAC 178

RESULT 5
BE259860
LOCUS      601154266F1 NIH_MGC_19 Homo sapiens cDNA clone IMAGE:3510113 5',
DEFINITION  mRNA sequence.
ACCESSION  BE259860
VERSION    BE259860.1 GI:9130682
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 476)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-re@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: LLCM194 row: d column: 18
High quality sequence stop: 475.
Location/Qualifiers
1..476
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3510113"
/tissue_type="neuroblastoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_19"
/Note="Organ: brain; Vector: pOTB7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAG(G). Library constructed by Ling Hong
in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZIP-cDNA synthesis kit
(Saratogene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."
BASE COUNT      110 a 146 c 139 g 81 t
ORIGIN

Query Match      100.0%; Score 18; DB 10; Length 476;
Best Local Similarity 100.0%; Pred. No. 4.4e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCTGAGGGAACAGTGAC 18
|||||
Db 182 TCTGAGGGAACAGTGAC 199

RESULT 6
CB149125
LOCUS      CB149125
DEFINITION  K-EST0205606 L15CKK1 Homo sapiens cDNA clone L15CKK1-35-D03 5',

```

```

mrna sequence.
ACCESSION CBI49125
VERSION CBI49125.1 GI:28131896
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 519)
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE 1 (bases 1 to 519)
JOURNAL Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
COMMENT Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
Kim,Y.S.
21C Frontier Korean EST Project 2001
Unpublished
Contact: Kim YS
Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Eoeun-Dong Yuseong-gu, Daejeon 305-333, South Korea
Tel.: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsung@mail.kribb.re.kr
Plate: 35 row: D column: 03
High quality sequence stop: 519.
Location/Qualifiers
1..519
/organism="Homo sapiens"
/mol_type="mrna"
/db_xref="taxon:9606"
/clone="L15CKK1-35-D03"
/sex="M"
/cell_lines="CK-K1"
/lab_host="Top10F"
/clone_lib="L15CKK1"
/note="Organ: Liver; Vector: pCNS-D2; Site:1: EcoRI;
Site:2: NotI; The poly (A)+ RNA was dephosphorylated with
bacterial alkaline phosphatase (BAP) and then decapped
with tobacco acid pyrophosphatase (TAP). The decapped
intact mRNA was ligated with DNA-RNA linker including
EcoRI site by treatment of T4 RNA ligase and the first
strand cDNA was synthesized from oligo dt-selected mRNA by
priming with dt-tailed vector. The dt-tailed vector was
adjusted to have about 60nt. The cDNA vector was
circularized with E. coli DNA ligase after digestion of
EcoRI which site is also included in vector. An RNA strand
converted to a DNA strand by Okayama-Berg method. The
obtained cDNA vectors were used for transformation of
competent cells E. coli Top10F by electroporation method.
The cDNA libraries constructed by this method are
full-length enriched cDNA library."
BASE COUNT 122 a 141 c 165 g 91 t
ORIGIN
Query Match 100.0%; Score 18; DB 14; Length 519;
Best Local Similarity 100.0%; Pred. No. 4.5e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCTGAGGGAACAGTGAC 18
|||||
Db 442 TCTGAGGGAACAGTGAC 459

RESULT 7
LOCUS BE386276 582 bp mRNA linear EST 21-JUL-2000
DEFINITION 601273495F1 NIH_MGC_20 Homo sapiens cDNA clone IMAGE:3614782 5',
mRNA sequence.
ACCESSION BE386276
VERSION BE386276.1 GI:9331641
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 582)
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE 1 (bases 1 to 582)
JOURNAL NIH-MGC http://mgc.nci.nih.gov/
COMMENT National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC/DCTD/DTP
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: LLCM286 row: j column: 06
High quality sequence stop: 612.
Location/Qualifiers
1..582
/organism="Homo sapiens"
/mol_type="mrna"
/db_xref="taxon:9606"
/clone="IMAGE:3614782"
/tissue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_20"
/note="Organ: skin; Vector: pOTB7; Site:1: XhoI; Site:2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
```

```

/db_xref="taxon:9606"
/clone="IMAGE:3618149"
/tissue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 20"
/notes="Organ: skin; Vector: pOTB7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGACAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
BASE COUNT      147 a 165 c 190 g 103 t
ORIGIN
Query Match      100.0%; Score 18; DB 10; Length 612;
Best Local Similarity 100.0%; Pred. No. 4.8e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCTGAGGGAACAGTGAC 18
|||||
587 TCTGAGGGAACAGTGAC 604

RESULT 9
AUI17078      639 bp mRNA linear EST 01-AUG-2002
LOCUS      AUI17078 HEMBA1 Homo sapiens cDNA clone HEMBA1000614 5', mRNA
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM      Homo sapiens (human)
REFERENCE
AUTHORS      Ota, T., Nishikawa, T., Suzuki, Y., Ishii, S., Saito, K., Kawai, Y.,
Yamamoto, J., Wakamatsu, A., Nakamura, Y., Nagai, T., Sugano, S. and
Isogai, T.
TITLE      HRI human cDNA project
JOURNAL
COMMENT
Contact: Takao Isogai
Genomics Laboratory
Helix Research Institute
1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: genomics@hri.co.jp
HRI human cDNA project; 5' - & 3' - end one pass sequencing: Helix
Research Institute; cDNA library construction: Department of
Virology, Institute of Medical Science, University of Tokyo, and
Helix Research Institute.
FEATURES
Location/Qualifiers
source
1..639
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="HEMBA1000614"
/tissue_type="whole embryo, mainly head"
/dev_stages="embryo, 10 weeks"
/clone_lib="HEMBA1"
/notes="Vector: pME18SFL3"
BASE COUNT      149 a 180 c 197 g 110 t
ORIGIN
Query Match      100.0%; Score 18; DB 9; Length 639;
Best Local Similarity 100.0%; Pred. No. 4.9e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCTGAGGGAACAGTGAC 18
|||||

```

```

Db 439 TCTGAGGGAACAGTGAC 456
RESULT 10
LOCUS      BG477932      644 bp mRNA linear EST 21-MAR-2001
DEFINITION      6025291f1 NIH_MGC_20 Homo sapiens cDNA clone IMAGE:4641386 5',
mRNA sequence.
ACCESSION      BG477932
VERSION      BG477932
KEYWORDS      EST.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
REFERENCE      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS      NIH-MGC http://mgs.nci.nih.gov/.
TITLE      National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL      Unpublished
COMMENT      Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC/DCTD/DTF
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM1408 row: e column: 03
High quality sequence stop: 639.
FEATURES
Location/Qualifiers
source
1..644
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4641386"
/tissue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 20"
/notes="Organ: skin; Vector: pOTB7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGACAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
BASE COUNT      151 a 183 c 196 g 112 t
ORIGIN
Query Match      100.0%; Score 18; DB 10; Length 644;
Best Local Similarity 100.0%; Pred. No. 4.9e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCTGAGGGAACAGTGAC 18
|||||
490 TCTGAGGGAACAGTGAC 507

RESULT 11
LOCUS      BE276291      682 bp mRNA linear EST 13-JUL-2000
DEFINITION      601144209f1 NIH_MGC_20 Homo sapiens cDNA clone IMAGE:3050019 5',
mRNA sequence.
ACCESSION      BE276291
VERSION      BE276291
KEYWORDS      EST.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
REFERENCE      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS      NIH-MGC http://mgs.nci.nih.gov/.

```

```

TITLE
JOURNAL
COMMENT
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC/DCTD/DTP
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: LLCM95 row: b column: 04
High quality sequence stop: 614.
Location/Qualifiers
1..682
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3050019"
/tissue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 20"
/notes="Organ: skin; Vector: pOTB7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
```

BASE COUNT 151 a 194 c 216 g 121 t

ORIGIN

```

Query Match 100.0%; Score 18; DB 10; Length 682;
Best Local Similarity 100.0%; Pred. No. 5e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCTGAGGGAACAGTGAC 18
|||||
Db 468 TCTGAGGGAACAGTGAC 485

RESULT 12
LOCUS BE895982 690 bp mRNA linear EST 20-OCT-2000
DEFINITION 601438641F1 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:3923602 5',
mRNA sequence.
ACCESSION BE895982
VERSION BE895982.1 GI:10359926
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NIH-MGC http://mgc.nci.nih.gov/
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC/DCTD/DTP
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM9760 row: a column: 11
High quality sequence stop: 582.
Location/Qualifiers
1..690
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"

FEATURES
source
```

```

/clone="IMAGE:3923602"
/tissue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 72"
/notes="Organ: skin; Vector: pCW-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 2 kb. Library constructed by Life
Technologies."
BASE COUNT 136 a 225 c 195 g 134 t
ORIGIN

Query Match 100.0%; Score 18; DB 10; Length 690;
Best Local Similarity 100.0%; Pred. No. 5e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCTGAGGGAACAGTGAC 18
|||||
Db 124 TCTGAGGGAACAGTGAC 141

RESULT 13
LOCUS BG476511 711 bp mRNA linear EST 21-MAR-2001
DEFINITION 602522049F1 NIH_MGC_20 Homo sapiens cDNA clone IMAGE:4640172 5',
mRNA sequence.
ACCESSION BG476511
VERSION BG476511.1 GI:13408790
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NIH-MGC http://mgc.nci.nih.gov/
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC/DCTD/DTP
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM405 row: b column: 13
High quality sequence stop: 711.
Location/Qualifiers
1..711
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4640172"
/tissue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 20"
/notes="Organ: skin; Vector: pOTB7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
```

BASE COUNT 173 a 189 c 225 g 123 t

ORIGIN

```

Query Match 100.0%; Score 18; DB 10; Length 711;
Best Local Similarity 100.0%; Pred. No. 5.1e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCTGAGGGAACAGTGAC 18
|||||
```



```

Db      573 TCTGAGGGAACAGTGAC 590

RESULT 14
LOCUS   BM042154
DEFINITION 603615791F1 NIH_MGC_112 Homo sapiens cDNA clone IMAGE:5420542 5',
mRNA sequence.
ACCESSION BM042154
VERSION   BM042154.1 GI:16771421
KEYWORDS  EST.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 725)
AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished
COMMENT   Contact: Robert Strausberg, Ph.D.
          Email: cgapbs-remail.nih.gov
          Tissue Procurement: DCTD/DTF
          cDNA Library Preparation: Ling Hong/Rubin Laboratory
          DNA Sequencing by: Incyte Genomics, Inc.
          Clone distribution: MGC clone distribution information can be
          found through the I.M.A.G.E. Consortium/LLNL at:
          http://image.llnl.gov
          Plate: LCM1875 row: e column: 23
          High quality sequence stop: 725.
FEATURES             source
    source            1..725
    Location/Qualifiers
        /organism="Homo sapiens"
        /mol_type="mRNA"
        /db_xref="taxon:9606"
        /clone="IMAGE:5420542"
        /tissue_type="melanotic melanoma, cell line"
        /lab_host="DH10B (phage-resistant)"
        /clone_lib="NIH_MGC_112"
        /note="Organ: skin; Vector: pORF7; Site 1: XhoI; Site 2:
        EcoRI; cDNA made by oligo-dT priming. Directionally cloned
        into EcoRI/XhoI sites using the following 5' adaptor:
        GGCACGAG(G). Library constructed by Ling Hong in the
        laboratory of Gerald M. Rubin (University of California,
        Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
        Superscript II RT (Life Technologies). Note: this is a
        NIH_MGC Library."
BASE COUNT      178 a _203 c 219 g 125 t
ORIGIN
Query Match      100.0%; Score 18; DB 12; Length 725;
Best Local Similarity 100.0%; Pred. No. 5.1e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 TCTGAGGGAACAGTGAC 18
        |||||
Db      502 TCTGAGGGAACAGTGAC 519

RESULT 15
LOCUS   BI771956
DEFINITION 603058981F1 NIH_MGC_122 Homo sapiens cDNA clone IMAGE:5208117 5',
mRNA sequence.
ACCESSION BI771956
VERSION   BI771956.1 GI:15763534
KEYWORDS  EST.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 761)
AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.

```

---

```

TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished
COMMENT   Contact: Robert Strausberg, Ph.D.
          Email: cgapbs-remail.nih.gov
          Tissue Procurement: Life Technologies, Inc.
          cDNA Library Preparation: Life Technologies, Inc.
          DNA Sequencing by: Incyte Genomics, Inc.
          Clone distribution: MGC clone distribution information can be
          found through the I.M.A.G.E. Consortium/LLNL at:
          http://image.llnl.gov
          Plate: LAM11522 row: b column: 22
          High quality sequence stop: 738.
FEATURES             source
    source            1..761
    Location/Qualifiers
        /organism="Homo sapiens"
        /mol_type="mRNA"
        /db_xref="taxon:9606"
        /clone="IMAGE:5208117"
        /lab_host="DH10B"
        /clone_lib="NIH_MGC_122"
        /note="Organ: pooled lung and spleen; Vector: pCMV-SPORT6;
        Site 1: NotI; Site 2: EcoRV (destroyed); RNA source
        anonymous pool of 24 week female lung, 16 week female
        spleen, and 20-22 week male spleens. Library is oligo-dT
        primed and directionally cloned (EcoRV site is destroyed
        upon cloning). Average insert size 1.4 kb, insert size
        range 1-3 kb. Library is normalized and enriched for
        full-length clones and was constructed by C. Gruber
        (Invitrogen). Research Genetics tracking code 026. Note:
        this is a NIH_MGC Library."
BASE COUNT      188 a 202 c 237 g 134 t
ORIGIN
Query Match      100.0%; Score 18; DB 12; Length 761;
Best Local Similarity 100.0%; Pred. No. 5.2e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 TCTGAGGGAACAGTGAC 18
        |||||
Db      573 TCTGAGGGAACAGTGAC 590

Search completed: August 19, 2003, 08:29:02
Job time : 59.6587 secs

```

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model  
Run on: August 18, 2003, 19:53:48 ; Search time 277.124 Seconds  
(without alignments)  
9743.047 Million cell updates/sec

Title: US-09-987-455-6  
Perfect score: 66  
Sequence: 1 atgaaagacagctatcgc.....ctaccgtggccagcgccg 66

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl.\*

- 1: gb\_ba.\*
- 2: gb\_htg.\*
- 3: gb\_in.\*
- 4: gb\_om.\*
- 5: gb\_ov.\*
- 6: gb\_pat.\*
- 7: gb\_ph.\*
- 8: gb\_pl.\*
- 9: gb\_pr.\*
- 10: gb\_ro.\*
- 11: gb\_sts.\*
- 12: gb\_sy.\*
- 13: gb\_un.\*
- 14: gb\_vi.\*
- 15: em\_ba.\*
- 16: em\_fun.\*
- 17: em\_hum.\*
- 18: em\_in.\*
- 19: em\_mu.\*
- 20: em\_om.\*
- 21: em\_or.\*
- 22: em\_ov.\*
- 23: em\_pat.\*
- 24: em\_ph.\*
- 25: em\_pl.\*
- 26: em\_ro.\*
- 27: em\_sts.\*
- 28: em\_un.\*
- 29: em\_vi.\*
- 30: em\_htg\_hum.\*
- 31: em\_htg\_inv.\*
- 32: em\_htg\_other.\*
- 33: em\_htg\_mus.\*
- 34: em\_htg\_pln.\*
- 35: em\_htg\_rpd.\*
- 36: em\_htg\_mam.\*
- 37: em\_htg\_vrt.\*
- 38: em\_sy.\*
- 39: em\_htgo\_hum.\*
- 40: em\_htgo\_mus.\*
- 41: em\_htgo\_other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	66	100.0	66	6	AX430829	AX430829 Sequence
2	66	100.0	66	6	AX431269	AX431269 Sequence
3	66	100.0	66	6	AX431272	AX431272 Sequence
4	66	100.0	705	12	HSIGKLC28	X95748 Synthetic C
5	66	100.0	705	12	HSIGKLC5	X95747 Synthetic C
6	66	100.0	711	12	HSIGKLC14	X95749 Synthetic C
7	66	100.0	711	12	HSIGKLC31	X95750 Synthetic C
8	66	100.0	1128	6	AX431268	AX431268 Sequence
9	66	100.0	1128	6	AX431271	AX431271 Sequence
10	66	100.0	3394	12	AF268280	AF268280 Phagemid
11	66	100.0	3450	6	AR241678	AR241678 Sequence
12	66	100.0	3758	12	AF268281	AF268281 Phagemid
13	66	100.0	5149	6	AX591634	AX591634 Sequence
14	66	100.0	5149	6	AX600065	AX600065 Sequence
15	66	100.0	5683	6	AX554424	AX554424 Sequence
16	66	100.0	6122	6	AX554413	AX554413 Sequence
17	66	100.0	6122	6	AX554422	AX554422 Sequence
18	66	100.0	6122	6	AX642149	AX642149 Sequence
19	66	100.0	6125	6	AX554420	AX554420 Sequence
20	64.4	97.6	5885	1	STAFRA	L42764 Staphylococ
21	61.2	92.7	580	6	AX742885	AX742885 Sequence
22	61.2	92.7	580	6	AX743509	AX743509 Sequence
23	60.2	91.2	74	6	E02814	E02814 DNA sequenc
24	60.2	91.2	74	6	E03955	E03955 DNA encodin
25	60.2	91.2	103	6	AX236721	AX236721 Sequence
26	60.2	91.2	249	6	I08200	I08200 Sequence 1
27	60.2	91.2	249	6	I52035	I52035 Sequence 1
28	60.2	91.2	438	12	SYNOMPANA	N62710 Synthetic 9
29	60.2	91.2	548	6	AR179433	AR179433 Sequence
30	60.2	91.2	548	6	AR282580	AR282580 Sequence
31	60.2	91.2	548	6	AX022441	AX022441 Sequence
32	60.2	91.2	548	6	AX030798	AX030798 Sequence
33	60.2	91.2	548	6	AX034619	AX034619 Sequence
34	60.2	91.2	652	6	E01685	E01685 hybrid DNA
35	60.2	91.2	739	1	AF234271	AF234271 Shigella
36	60.2	91.2	746	1	AF234270	AF234270 Salmonell
37	60.2	91.2	1400	1	STOMPA	X02006 Salmonella
38	60.2	91.2	1697	1	AF234269	AF234269 Escherich
39	60.2	91.2	2270	6	AX191724	AX191724 Sequence
40	60.2	91.2	2271	1	ECOMPA	V00307 E. coli omp
41	60.2	91.2	6477	6	A08458	A08458 pinf 4-49 m
42	60.2	91.2	7499	12	ASAK3121	AJ223121 Artificia
43	60.2	91.2	7500	12	ASAK3122	AJ223122 Artificia
C 44	60.2	91.2	11639	1	AE000198	AE000198 Escherich
C 45	60.2	91.2	11641	1	AE005286	AE005286 Escherich

ALIGNMENTS

RESULT 1  
AX430829  
LOCUS  
DEFINITION Sequence 1 from Patent WO0240696.  
ACCESSION AX430829  
VERSION AX430829.1 GI:21655908  
KEYWORDS Escherichia coli  
SOURCE Escherichia coli  
ORGANISM Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
Enterobacteriaceae; Escherichia.  
REFERENCE  
AUTHORS Goetz,F., Werner,R.G., Tayapiwatana,C., Manosroi,A. and Manosroi,J.  
TITLE Methods for large scale protein production in prokaryotes  
JOURNAL Patent: WO 0240696-A 1 23-MAY-2002;

```

FEATURES             BOEHRINGER INGELHEIM INT (DE)
source               Location/Qualifiers
                     1..66
                     /organism="Escherichia coli"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:562"
BASE COUNT          14 a 18 c 21 g 13 t
ORIGIN
Query Match          100.0%; Score 66; DB 6; Length 66;
Best Local Similarity 100.0%; Pred. No. 1.4e-11;
Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATGAAAAGACAGCTATCGCGATTGCGAGTGCAGTGGCTGTTTCGCTACCGTGCGCCAG 60
Db 1 ATGAAAAGACAGCTATCGCGATTGCGAGTGCAGTGGCTGTTTCGCTACCGTGCGCCAG 60
Qy 61 GCGGCC 66
Db 61 GCGGCC 66

RESULT 2
AX431269             AX431269             66 bp      DNA      linear      PAT 28-JUN-2002
LOCUS                Sequence 3 from Patent WO0240650.
ACCESSION            AX431269
VERSION              AX431269.1 GI:21656151
KEYWORDS              Escherichia coli
SOURCE                Escherichia coli
ORGANISM              Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
                      Enterobacteriaceae; Escherichia.
REFERENCE
AUTHORS              Goetz,F., Werner,R.G., Tayapiwatana,C., Manosroi,A. and Manosroi,J.
TITLE                Methods for large scale production of recombinant dna-derived tpa
                      or k2s molecules
JOURNAL              Patent: WO 0240650-A 3 23-MAY-2002;
                      BOEHRINGER INGELHEIM INT (DE)
FEATURES
source               Location/Qualifiers
                     1..66
                     /organism="Escherichia coli"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:562"
BASE COUNT          14 a 18 c 21 g 13 t
ORIGIN
Query Match          100.0%; Score 66; DB 6; Length 66;
Best Local Similarity 100.0%; Pred. No. 1.4e-11;
Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATGAAAAGACAGCTATCGCGATTGCGAGTGCAGTGGCTGTTTCGCTACCGTGCGCCAG 60
Db 1 ATGAAAAGACAGCTATCGCGATTGCGAGTGCAGTGGCTGTTTCGCTACCGTGCGCCAG 60
Qy 61 GCGGCC 66
Db 61 GCGGCC 66

RESULT 3
AX431272             AX431272             66 bp      DNA      linear      PAT 28-JUN-2002
LOCUS                Sequence 6 from Patent WO0240650.
ACCESSION            AX431272
VERSION              AX431272.1 GI:21656154
KEYWORDS              Escherichia coli
SOURCE                Escherichia coli
ORGANISM              Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
                      Enterobacteriaceae; Escherichia.
REFERENCE
AUTHORS              Goetz,F., Werner,R.G., Tayapiwatana,C., Manosroi,A. and Manosroi,J.
TITLE                Methods for large scale production of recombinant dna-derived tpa
                      or k2s molecules
JOURNAL              Patent: WO 0240650-A 3 23-MAY-2002;
                      BOEHRINGER INGELHEIM INT (DE)
FEATURES
source               Location/Qualifiers
                     1..66
                     /organism="Escherichia coli"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:562"
BASE COUNT          14 a 18 c 21 g 13 t
ORIGIN
Query Match          100.0%; Score 66; DB 6; Length 66;
Best Local Similarity 100.0%; Pred. No. 1.4e-11;
Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATGAAAAGACAGCTATCGCGATTGCGAGTGCAGTGGCTGTTTCGCTACCGTGCGCCAG 60
Db 1 ATGAAAAGACAGCTATCGCGATTGCGAGTGCAGTGGCTGTTTCGCTACCGTGCGCCAG 60
Qy 61 GCGGCC 66
Db 61 GCGGCC 66

TITLE               Methods for large scale production of recombinant dna-derived tpa
or k2s molecules
JOURNAL             Patent: WO 0240650-A 3 23-MAY-2002;
BOEHRINGER INGELHEIM INT (DE)
FEATURES
source             Location/Qualifiers
1..66
/organism="Escherichia coli"
/mol_type="genomic DNA"
/db_xref="taxon:562"
BASE COUNT        14 a 18 c 21 g 13 t
ORIGIN
Query Match        100.0%; Score 66; DB 6; Length 66;
Best Local Similarity 100.0%; Pred. No. 1.4e-11;
Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATGAAAAGACAGCTATCGCGATTGCGAGTGCAGTGGCTGTTTCGCTACCGTGCGCCAG 60
Db 1 ATGAAAAGACAGCTATCGCGATTGCGAGTGCAGTGGCTGTTTCGCTACCGTGCGCCAG 60
Qy 61 GCGGCC 66
Db 61 GCGGCC 66

RESULT 4
HSIGKLC28           HSIGKLC28           705 bp      mRNA      linear      SYN 29-MAR-2001
LOCUS               Synthetic construct including Homo sapiens immunoglobulin kappa
DEFINITION          chain (clone: 28).
ACCESSION           X95748
VERSION             X95748.1 GI:1514578
KEYWORDS             constant region; immunoglobulin; kappa light chain.
SOURCE              synthetic construct
ORGANISM             artificial sequences.
REFERENCE
1                   Steinberger,P., Kraft,D. and Valenta,R.
AUTHORS              Construction of a combinatorial IGE library from an allergic
TITLE                patient. Isolation and characterization of human IGE Fabs with
                      specificity for the major timothy grass pollen allergen, Phl p 5
JOURNAL              J. Biol. Chem. 271 (18), 10967-10972 (1996)
MEDLINE             8631038
PUBMED              8631916
REFERENCE            2 (bases 1 to 705)
AUTHORS              Valenta,R.L.S.
TITLE                Direct Submission
JOURNAL              Submitted (19-FEB-1996) R.L.S. Valenta, Institute of General &
                      Experimental Pathology, General Hospital, Waehringer Guertel 18-20,
                      1090 Vienna, AUSTRIA
FEATURES
source             Location/Qualifiers
1..705
/organism="synthetic construct"
/mol_type="mRNA"
/db_xref="taxon:32630"
/focus
1..66
/organism="Escherichia coli"
/mol_type="mRNA"
/db_xref="taxon:562"
67..705
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="28"
/cell_type="B-cell"
/clone_lib="pcomb3H"
1..705
/gene="OMP-A"
1..705
/gene="OMP-A"
/codon_start=1
/transl_table=11

```





TITLE Construction of a combinatorial IgE library from an allergic patient. Isolation and characterization of human IgE Fabs with specificity for the major timothy grass pollen allergen, Phl p 5

JOURNAL J. Biol. Chem. 271 (18), 10967-10972 (1996)

MEDLINE 96210038

PUBMED 8631916

REFERENCE 2 (bases 1 to 711)

AUTHORS Valenta, R.L.S.

TITLE Direct Submission

JOURNAL Submitted (19-FEB-1996) R.L.S. Valenta, Institute of General & Experimental Pathology, General Hospital, Waehringer Guertel 18-20, 1090 Vienna, AUSTRIA

FEATURES

source

1..711

/organism="synthetic construct"

/mol\_type="mRNA"

/db\_xref="taxon:32630"

/focus

1..66

/organism="Escherichia coli"

/mol\_type="mRNA"

/db\_xref="taxon:562"

67..711

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clones="31"

/cell\_type="B-cell"

/clone\_lib="pcomb3H"

1..711

/gene="OMP-A"

1..711

/gene="OMP-A"

/codon\_start=1

/transl\_table=11

/product="immunoglobulin kappa light chain"

/protein\_id="CAA65061.1"

/db\_xref="GI:1514583"

/db\_xref="REMBL:CAA65061"

/translation="MKTKAIAIAVALAGFATVAQAEIVMTQSPSSLSAFVGDRTVITCRASQISVLYNMVQKPGKAPKLLIYAASLOSQVSPSPSGSGTDFLTITSSLOPEDFATYYCQSHSTPYPFGQGTNLEIKRTVAAPSVPFPPSPDEOLKSGTASVUCLLN FYPREAKVQKVDNALQSGNSQESVTEQDSKDSYSTISLTLSKADYKHKVYACV THQGLSPFVTKSFNRGEC"

mRNA

1..66

/gene="OMP-A"

/note="leaderpeptide gene"

67..135

/gene="OMP-A"

/note="FR1"

67..84

/gene="OMP-A"

/note="PCR-primer sequence"

67..72

/gene="OMP-A"

/note="SAC 1 restriction site"

136..168

/gene="OMP-A"

/note="CDR1"

169..213

/gene="OMP-A"

/note="FR2"

214..234

/gene="OMP-A"

/note="CDR2"

235..330

/gene="OMP-A"

/note="FR3"

331..357

/gene="OMP-A"

/note="CDR3"

358..390

/gene="OMP-A"

C\_region

/note="FR4"

391..708

/gene="OMP-A"

/note="kappa constant region"

667..708

/gene="OMP-A"

/note="PCR-primer sequence"

188 a 194 c 179 g 150 t

BASE COUNT

ORIGIN

Query Match 100.0%; Score 66; DB 12; Length 711;

Best Local Similarity 100.0%; Pred. No. 1.6e-11;

Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGAAAAAGACAGCTATCGCGATTGCGACTGGCGACTGGCTGGTTCGCTACCGTGCCCGAG 60

|||||

Db 1 ATGAAAAAGACAGCTATCGCGATTGCGACTGGCGACTGGCTGGTTCGCTACCGTGCCCGAG 60

Qy 61 GCGGCC 66

|||||

Db 61 GCGGCC 66

RESULT 8

AX431268

LOCUS AX431268 1128 bp DNA linear PAT 28-JUN-2002

DEFINITION Sequence 2 from Patent WO0240650.

ACCESSION AX431268

VERSION AX431268.1 GI:21656150

KEYWORDS synthetic construct

SOURCE synthetic construct

ORGANISM artificial sequences.

REFERENCE 1

AUTHORS Goetz, F., Werner, R.G., Tayapiwatana, C., Manosroi, A. and Manosroi, J.

TITLE Methods for large scale production of recombinant dna-derived tpa or k2a molecules

JOURNAL Patent: WO 0240650-A 2 23-MAY-2002;

BOEHRINGER INGELHEIM INT (DE)

FEATURES

Location/Qualifiers

1..1128

/organism="synthetic construct"

/mol\_type="genomic DNA"

/db\_xref="taxon:32630"

/note="coding sequence for OmpA-K2S fusion protein"

238 a 332 c 332 g 226 t

BASE COUNT

ORIGIN

Query Match 100.0%; Score 66; DB 6; Length 1128;

Best Local Similarity 100.0%; Pred. No. 1.6e-11;

Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGAAAAAGACAGCTATCGCGATTGCGACTGGCGACTGGCTGGTTCGCTACCGTGCCCGAG 60

|||||

Db 1 ATGAAAAAGACAGCTATCGCGATTGCGACTGGCGACTGGCTGGTTCGCTACCGTGCCCGAG 60

Qy 61 GCGGCC 66

|||||

Db 61 GCGGCC 66

RESULT 9

AX431271

LOCUS AX431271 1128 bp DNA linear PAT 28-JUN-2002

DEFINITION Sequence 5 from Patent WO0240650.

ACCESSION AX431271

VERSION AX431271.1 GI:21656153

KEYWORDS synthetic construct

SOURCE synthetic construct

ORGANISM artificial sequences.

REFERENCE 1

AUTHORS Goetz, F., Werner, R.G., Tayapiwatana, C., Manosroi, A. and Manosroi, J.

TITLE Methods for large scale production of recombinant dna-derived tpa  
or k2s molecules  
JOURNAL Patent: WO 0240650-A 5 23-MAY-2002;  
BOEHRINGER INGELHEIM INT (DE)  
FEATURES Location/Qualifiers  
source 1. 1128  
/organism="synthetic construct"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:32630"  
/note="coding sequence for OmpA-K2S fusion protein"  
BASE COUNT 238 a 332 c 332 g 226 t  
ORIGIN  
Query Match 100.0%; Score 66; DB 6; Length 1128;  
Best Local Similarity 100.0%; Pred. No. 1.7e-11; Length 1128;  
Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 ATGAAAAGACAGCTATCGGATTGCGATGCGACTGGCTGTTTCGTACCGTGGCCAG 60  
Db 1 ATGAAAAGACAGCTATCGGATTGCGATGCGACTGGCTGTTTCGTACCGTGGCCAG 60  
Qy 61 GCGGCC 66  
Db 61 GCGGCC 66  
RESULT 10  
AF268280 3394 bp DNA circular SYN 03-OCT-2000  
LOCUS Phagemid cloning vector pComb3H, complete sequence.  
DEFINITION AF268280  
ACCESSION AF268280.1 GI:10505048  
VERSION  
KEYWORDS Phagemid cloning vector pComb3H  
ORGANISM Phagemid cloning vector pComb3H  
SOURCE Phagemid cloning vector pComb3H  
REFERENCE 1 (bases 1 to 3394)  
AUTHORS Rader, C. and Barbas, C.F. III.  
JOURNAL (in) PHAGE DISPLAY, A LABORATORY MANUAL, Cold Spring Harbor  
Laboratory Press, Cold Spring Harbor, NY, USA (2000), In press  
REFERENCE 2 (bases 1 to 3394)  
AUTHORS Rader, C. and Barbas, C.F. III.  
JOURNAL Direct Submission  
TITLE Submitted (16-MAY-2000) Department of Molecular Biology, BCC-526,  
Scripps Research Institute, 10550 N. Torrey Pines Rd., La Jolla, CA  
92037, USA  
FEATURES Location/Qualifiers  
source 1. 3394  
/organism="Phagemid cloning vector pComb3H"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:137785"  
BASE COUNT 862 a 783 c 876 g 873 t  
ORIGIN  
Query Match 100.0%; Score 66; DB 12; Length 3394;  
Best Local Similarity 100.0%; Pred. No. 1.7e-11;  
Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 ATGAAAAGACAGCTATCGGATTGCGATGCGACTGGCTGTTTCGTACCGTGGCCAG 60  
Db 222 ATGAAAAGACAGCTATCGGATTGCGATGCGACTGGCTGTTTCGTACCGTGGCCAG 281  
Qy 61 GCGGCC 66  
Db 282 GCGGCC 287  
RESULT 11  
AR241678 3450 bp DNA linear PAT 20-DEC-2002  
LOCUS Sequence 20 from patent US 6472147.  
DEFINITION AR241678  
ACCESSION AR241678.1 GI:27287480  
VERSION

KEYWORDS Unknown.  
SOURCE Unknown.  
ORGANISM Unclassified.  
REFERENCE 1 (bases 1 to 3450)  
AUTHORS Janda, K.D., Wirsching, P., Lerner, R.A. and Gao, C.  
TITLE Methods for display of heterodimeric proteins on filamentous phage  
using pVII and pIX, compositions, vectors and combinatorial  
libraries  
JOURNAL Patent: US 6472147-A 20 29-OCT-2002;  
FEATURES Location/Qualifiers  
source 1. 3450  
/organism="unknown"  
BASE COUNT 873 a 799 c 891 g 887 t  
ORIGIN  
Query Match 100.0%; Score 66; DB 6; Length 3450;  
Best Local Similarity 100.0%; Pred. No. 1.7e-11;  
Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 ATGAAAAGACAGCTATCGGATTGCGATGCGACTGGCTGTTTCGTACCGTGGCCAG 60  
Db 2667 ATGAAAAGACAGCTATCGGATTGCGATGCGACTGGCTGTTTCGTACCGTGGCCAG 2726  
Qy 61 GCGGCC 66  
Db 2727 GCGGCC 2732  
RESULT 12  
AF268281 3758 bp DNA circular SYN 03-OCT-2000  
LOCUS Phagemid cloning vector pComb3X, complete sequence.  
DEFINITION AF268281  
ACCESSION AF268281  
VERSION AF268281.1 GI:10505049  
KEYWORDS Phagemid cloning vector pComb3X  
ORGANISM Phagemid cloning vector pComb3X  
SOURCE Phagemid cloning vector pComb3X  
REFERENCE 1 (bases 1 to 3758)  
AUTHORS Rader, C. and Barbas, C.F. III.  
JOURNAL (in) PHAGE DISPLAY, A LABORATORY MANUAL, Cold Spring Harbor  
Laboratory Press, Cold Spring Harbor, NY, USA (2000), In press  
REFERENCE 2 (bases 1 to 3758)  
AUTHORS Rader, C. and Barbas, C.F. III.  
JOURNAL Direct Submission  
TITLE Submitted (16-MAY-2000) Department of Molecular Biology, BCC-526,  
Scripps Research Institute, 10550 N. Torrey Pines Rd., La Jolla, CA  
92037, USA  
FEATURES Location/Qualifiers  
source 1. 3758  
/organism="Phagemid cloning vector pComb3X"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:137786"  
BASE COUNT 939 a 896 c 959 g 964 t  
ORIGIN  
Query Match 100.0%; Score 66; DB 12; Length 3758;  
Best Local Similarity 100.0%; Pred. No. 1.7e-11;  
Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 ATGAAAAGACAGCTATCGGATTGCGATGCGACTGGCTGTTTCGTACCGTGGCCAG 60  
Db 222 ATGAAAAGACAGCTATCGGATTGCGATGCGACTGGCTGTTTCGTACCGTGGCCAG 281  
Qy 61 GCGGCC 66  
Db 282 GCGGCC 287  
RESULT 13  
AX591634 5149 bp DNA linear PAT 27-JAN-2003  
LOCUS AX591634



DEFINITION Sequence 8 from Patent WO0246434.

ACCESSION AX591634.1 GI:27950030

VERSION AX591634.1 GI:27950030

KEYWORDS synthetic construct

SOURCE synthetic construct

ORGANISM synthetic construct

artificial sequences.

REFERENCE 1

AUTHORS Bowdish,K.S. and Barbas-Frederickson,S.

TITLE Novel plasmid vectors

JOURNAL Patent: WO 0246434-A 8 13-JUN-2002;

ALEXION PHARMACEUTICALS, INC. (US)

FEATURES Location/Qualifiers

source

1..5149

/organism="synthetic construct"

/mol\_type="genomic DNA"

/db\_xref="taxon:32630"

/note="vector"

BASE COUNT 1279 a 1244 c 1372 g 1254 t

ORIGIN

Query Match 100.0%; Score 66; DB 6; Length 5149;

Best Local Similarity 100.0%; Pred. No. 1.8e-11;

Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 ATGAAAAGACAGCTATCGGATTCGAGTGCAGTGGCTGGTTTCGCTACCGTGCCCCAG 60

Db 2611 ATGAAAAGACAGCTATCGGATTCGAGTGCAGTGGCTGGTTTCGCTACCGTGCCCCAG 2670

Oy 61 GCGGCC 66

Db 2671 GCGGCC 2676

RESULT 14

AX600065

LOCUS

AX600065

DEFINITION Sequence 60 from Patent WO0246238.

ACCESSION AX600065

VERSION AX600065.1 GI:28400138

KEYWORDS synthetic construct

SOURCE synthetic construct

ORGANISM synthetic construct

artificial sequences.

REFERENCE 1

AUTHORS Bowdish,K.S., Barbas-Frederickson,S. and Renshaw,M.

TITLE Rationally designed antibodies

JOURNAL Patent: WO 0246238-A 60 13-JUN-2002;

ALEXION PHARMACEUTICALS, INC. (US)

FEATURES Location/Qualifiers

source

1..5149

/organism="synthetic construct"

/mol\_type="genomic DNA"

/db\_xref="taxon:32630"

/note="vector"

BASE COUNT 1279 a 1244 c 1372 g 1254 t

ORIGIN

Query Match 100.0%; Score 66; DB 6; Length 5149;

Best Local Similarity 100.0%; Pred. No. 1.8e-11;

Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 ATGAAAAGACAGCTATCGGATTCGAGTGCAGTGGCTGGTTTCGCTACCGTGCCCCAG 60

Db 2611 ATGAAAAGACAGCTATCGGATTCGAGTGCAGTGGCTGGTTTCGCTACCGTGCCCCAG 2670

Oy 61 GCGGCC 66

Db 2671 GCGGCC 2676

RESULT 15

AX554424

LOCUS

AX554424

DEFINITION Sequence 12 from Patent WO0246436.

ACCESSION AX554424

VERSION AX554424.1 GI:25898200

KEYWORDS synthetic construct

SOURCE synthetic construct

ORGANISM synthetic construct

artificial sequences.

REFERENCE 1

AUTHORS Bowdish,K.S., Barbas-Frederickson,S., Wild,M. and McWhirter,J.

TITLE Novel plasmid vectors

JOURNAL Patent: WO 0246436-A 12 13-JUN-2002;

ALEXION PHARMACEUTICALS, INC. (US)

FEATURES Location/Qualifiers

source

1..5683

/organism="synthetic construct"

/mol\_type="genomic DNA"

/db\_xref="taxon:32630"

/note="plasmid"

BASE COUNT 1337 a 1492 c 1543 g 1311 t

ORIGIN

Query Match 100.0%; Score 66; DB 6; Length 5683;

Best Local Similarity 100.0%; Pred. No. 1.8e-11;

Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 ATGAAAAGACAGCTATCGGATTCGAGTGCAGTGGCTGGTTTCGCTACCGTGCCCCAG 60

Db 2611 ATGAAAAGACAGCTATCGGATTCGAGTGCAGTGGCTGGTTTCGCTACCGTGCCCCAG 2670

Oy 61 GCGGCC 66

Db 2671 GCGGCC 2676

Search completed: August 19, 2003, 04:46:35

Job time : 278.124 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 18, 2003, 19:51:53 ; Search time 23.2513 Seconds  
(without alignments)  
7662.485 Million cell updates/sec

Title: US-09-987-455-6

Perfect score: 66

Sequence: 1 atgaaaagacagtatcgc.....ctaccgtggccaggcgcc 66

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N\_Geneseq\_19Jun03.\*

1: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.\*  
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.\*  
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.\*  
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.\*  
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.\*  
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.\*  
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.\*  
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.\*  
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.\*  
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.\*  
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.\*  
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.\*  
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.\*  
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.\*  
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.\*  
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.\*  
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.\*  
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.\*  
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.\*  
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.\*  
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.\*  
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.\*  
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.\*  
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.\*  
25: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	66	100.0	66	AAD40615	Escherichia coli o
2	66	100.0	66	AAD38987	Escherichia coli O
3	66	100.0	1128	AAD40614	OmpA-K2S fusion pr
4	66	100.0	1574	AAZ57599	Erythrocyte and He
5	66	100.0	3450	AAZ66993	Vector pCGMT-1b.
6	66	100.0	4691	AAQ52546	pcom3 expression
7	66	100.0	5149	ABQ73372	Plasmid pRL8 nucle
8	66	100.0	5149	ABN84077	Plasmid vector pr

9	66	100.0	5683	24	AAL41116	Plasmid pRL5 nucle
10	66	100.0	6122	24	AAL41112	Plasmid pRL5 CAT n
11	66	100.0	6122	24	AAL41115	Plasmid pRL5 bsi-C
12	66	100.0	6122	24	ABN84078	Plasmid pRL5-CAT.
13	66	100.0	6125	24	AAL41114	Plasmid pRL5 CAT-A
14	66	100.0	6166	16	AAQ92547	Expression vector;
15	60.2	91.2	73	12	AAQ10569	Omp A signal pepti
16	60.2	91.2	103	22	AAQ44154	Neisseria meningit
17	60.2	91.2	249	11	AAQ03521	TAC recombinant DN
18	60.2	91.2	548	18	AAQ91586	Growth factor LHL
19	60.2	91.2	548	20	AAQ34590	DNA encoding a LHL
20	60.2	91.2	652	9	AAQ81344	Sequence encoding
21	60.2	91.2	2814	14	AAQ41731	PE binding/translo
22	60.2	91.2	2814	14	AAQ38414	PE binding and tra
23	60.2	91.2	6477	11	AAQ02030	Plasmid pinf 4-49.
24	59.6	90.3	73	21	AAA12579	Synthetic DNA enco
25	59.6	90.3	415	17	AAQ15982	pSKAN8 fragment #1
26	59.6	90.3	470	18	AAQ91590	Kappa light chain
27	59.6	90.3	470	20	AAQ34594	DNA encoding a hum
28	59.6	90.3	599	18	AAQ91588	Growth factor LHL.
29	59.6	90.3	599	20	AAQ34593	The ORF of the clo
30	59.6	90.3	759	19	AAQ53721	Cloned duplicate o
31	59.6	90.3	770	19	AAQ53720	Plasmid pBP22 DNA
32	59.6	90.3	793	20	AAQ24101	Billin binding-prot
33	59.6	90.3	793	22	AAQ25703	Qm212 single chain
34	59.6	90.3	932	22	AAA91151	Expression cassett
35	59.6	90.3	932	22	AAA91151	Growth factor TLHL
36	59.6	90.3	1031	18	AAQ91589	DNA encoding TLHL.
37	59.6	90.3	1031	20	AAQ34592	Plasmid pBP20 DNA
38	59.6	90.3	1219	20	AAQ24099	Billin binding-prot
39	59.6	90.3	1219	22	AAQ25695	Billin binding-prot
40	59.6	90.3	1219	22	AAQ25708	Plasmid pBP21 DNA
41	59.6	90.3	1380	20	AAQ24100	Billin binding-prot
42	59.6	90.3	1380	22	AAQ25710	DNA encoding a ccm
43	59.6	90.3	1479	20	AAQ34596	Growth factor CATB
44	59.6	90.3	1490	18	AAQ91587	DNA encoding CATAB
45	59.6	90.3	1490	20	AAQ34591	

#### ALIGNMENTS

##### RESULT 1

AAD40615  
ID AAD40615 standard; DNA; 66 BP.

XX  
AC AAD40615;

DT 30-OCT-2002 (first entry)

DE Escherichia coli ompA peptide encoding DNA.

XX Tissue plasminogen activator; tPA; kringle 2 serine protease; stroke;  
KW K2S; cardiac infarction; acute myocardial infarction; pulmonary embolism;  
KW artery occlusion; deep vein thrombosis; blood clotting; thrombolytic;  
KW cerebroprotective; cardiant; ompA; gene; ds.

OS Escherichia coli.

XX Key Location/Qualifiers

FT CDS 1..66

FT /\*tag= a

FT /product= "OmpA peptide"

FT /note= "No stop codon"

FT /partial

XX WO200240650-A2.

PN 23-MAY-2002.

PD 07-NOV-2001; 2001WO-EPI2857.

XX 14-NOV-2000; 2000GE-0027779.

PR

XX (BOEH ) BOEHRINGER INGELHEIM INT GMBH.  
 XX Werner R, Goetz F, Tayapiwatana C, Manosroi J, Manosroi A;  
 XX WPI; 2002-519376/55.  
 XX P-PSDB; AAE25035.  
 XX  
 XX Producing active, correctly folded recombinant tissue plasminogen  
 PT activator, Kringle 2 serine protease in prokaryotic cells by expressing  
 PT the protein-encoding DNA operably linked to DNA coding for signal  
 XX peptide OmpA -  
 XX  
 PS Claim 8; Page 30; 80pp; English.  
 XX  
 CC The present invention relates to a method of producing extracellularly  
 CC secreted, active, correctly folded, recombinant tissue plasminogen  
 CC activator (tPA), Kringle 2 serine protease molecule (K2S) or their  
 CC variants in prokaryotic cells by expressing the protein-encoding DNA  
 CC operably linked to DNA coding for signal peptide OmpA. The method is  
 CC useful for producing recombinant DNA-derived tPA, K2S or their variants.  
 CC Sequences of the invention are useful for manufacturing a medicament  
 CC for treating stroke, cardiac infarction, acute myocardial infarction,  
 CC pulmonary-embolism, any artery occlusion such as intracranial artery  
 CC occlusion (e.g., arteries supplying the brain), peripherally occluded  
 CC arteries, coronary artery occlusion, deep vein thrombosis or related  
 CC diseases associated with unwanted blood clotting. The present sequence  
 CC is a DNA fragment encoding Escherichia coli OmpA peptide.  
 XX  
 SQ Sequence 66 BP; 14 A; 18 C; 21 G; 13 T; 0 other;  
 Query Match 100.0%; Score 66; DB 24; Length 66;  
 Best Local Similarity 100.0%; Pred. No. 2.1e-14;  
 Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 ATGAAAAGACAGCTATCGCGATTGCGAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCGAG 60  
 Db 1 ATGAAAAGACAGCTATCGCGATTGCGAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCGAG 60  
 Qy 61 GCGGCC 66  
 Db 61 GCGGCC 66  
 RESULT 2  
 AAD38987  
 ID AAD38987 standard; DNA; 66 BP.  
 AC AAD38987;  
 DT 23-SEP-2002 (first entry)  
 XX Escherichia coli OmpA DNA.  
 XX Recombinant DNA-derived heterologous protein; OmpA; K2S; kringle 2;  
 KW serine protease; ds.  
 XX Escherichia coli.  
 OS WO200240696-A2.  
 XX  
 PN WO200240696-A2.  
 XX  
 PD 23-MAY-2002.  
 XX  
 XX 08-NOV-2001; 2001WO-EP12920.  
 XX  
 XX 14-NOV-2000; 2000GB-0027782.  
 XX  
 XX (BOEH ) BOEHRINGER INGELHEIM INT GMBH.  
 XX Werner R, Goetz F, Tayapiwatana C, Manosroi J, Manosroi A;  
 XX WPI; 2002-471625/50.  
 XX  
 PT Producing recombinant DNA-derived kringle 2 plus serine protease,  
 PT comprises using a prokaryotic cell expressing a vector having a DNA  
 PT coding for a heterologous protein operably linked to a DNA coding for  
 XX the signal peptide OmpA -  
 XX  
 PS Claim 9; Page 23; 52pp; English.  
 XX  
 CC The invention relates to a method for producing recombinant DNA-derived  
 CC heterologous protein in prokaryotic cells, where the heterologous protein  
 CC is secreted extracellularly as an active and correctly folded protein and  
 CC the prokaryotic cell contains and expresses a vector comprising the DNA  
 CC coding for the heterologous protein operably linked to the DNA coding  
 CC for the signal peptide OmpA or its functional derivative. The method is  
 CC useful for commercial large-scale production of heterologous proteins,  
 CC e.g. K2S (kringle 2 plus serine protease), in prokaryotic cells, and is  
 CC generally applicable in the expression of several different proteins and  
 CC polypeptides which do not require mammalian glycosylation in prokaryotic  
 CC host cells. The method may also be used to obtain DNA sequences of a  
 CC protein of interest to be expressed from databases and cloned for use.  
 CC The present sequence is Escherichia coli OmpA DNA.  
 XX  
 SQ Sequence 66 BP; 14 A; 18 C; 21 G; 13 T; 0 other;  
 Query Match 100.0%; Score 66; DB 24; Length 66;  
 Best Local Similarity 100.0%; Pred. No. 2.1e-14;  
 Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 ATGAAAAGACAGCTATCGCGATTGCGAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCGAG 60  
 Db 1 ATGAAAAGACAGCTATCGCGATTGCGAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCGAG 60  
 Qy 61 GCGGCC 66  
 Db 61 GCGGCC 66  
 RESULT 3  
 AAD40614  
 ID AAD40614 standard; DNA; 1128 BP.  
 XX  
 AC AAD40614;  
 DT 30-OCT-2002 (first entry)  
 XX OmpA-K2S fusion protein encoding DNA.  
 XX Tissue plasminogen activator; tPA; kringle 2 serine protease; stroke;  
 KW K2S; cardiac infarction; acute myocardial infarction; pulmonary embolism;  
 KW artery occlusion; deep vein thrombosis; blood clotting; thrombolytic;  
 KW cerebroprotective; cardiant; ompA; fusion protein; gene; ds.  
 XX  
 OS Unidentified.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 1..1128  
 FT /\*tag= a  
 FT /product= "OmpA-K2S fusion protein"  
 FT /transl\_except= (pos:1126..1128, aa:Pro-Gly)  
 FT /note= "No stop codon"  
 FT /partial  
 XX  
 PN WO200240650-A2.  
 XX  
 PD 23-MAY-2002.  
 XX  
 XX 07-NOV-2001; 2001WO-EP12857.  
 XX  
 XX 14-NOV-2000; 2000GB-0027779.  
 XX  
 XX (BOEH ) BOEHRINGER INGELHEIM INT GMBH.  
 XX Werner R, Goetz F, Tayapiwatana C, Manosroi J, Manosroi A;  
 XX WPI; 2002-471625/50.  
 XX

DR WPI; 2002-519376/55.  
 DR P-PSDB; AAE25034.  
 PT Producing active, correctly folded recombinant tissue plasminogen  
 PT activator, Kringle 2 serine protease in prokaryotic cells by expressing  
 PT the protein-encoding DNA operably linked to DNA coding for signal  
 PT peptide OmpA -  
 XX  
 XX Claim 7; Page 30; 80pp; English.  
 PS  
 XX The present invention relates to a method of producing extracellularly  
 CC secreted, active, correctly folded, recombinant tissue plasminogen  
 CC activator (tPA), Kringle 2 serine protease molecule (K2S) or their  
 CC variants in prokaryotic cells by expressing the protein-encoding DNA  
 CC operably linked to DNA coding for signal peptide OmpA. The method is  
 CC useful for producing recombinant DNA-derived tPA, K2S or their variants.  
 CC Sequences of the invention are useful for manufacturing a medicament  
 CC for treating stroke, cardiac infarction, acute myocardial infarction,  
 CC pulmonary embolism, any artery occlusion such as intracranial artery  
 CC occlusion (e.g., arteries supplying the brain), peripherally occluded  
 CC arteries, coronary artery occlusion, deep vein thrombosis or related  
 CC diseases associated with unwanted blood clotting. The present sequence  
 CC is a DNA encoding a fusion protein comprising OmpA and K2S protein.  
 XX  
 XX Sequence 1128 BP; 238 A; 332 C; 332 G; 226 T; 0 other;  
 SQ

Query Match 100.0%; Score 66; DB 24; Length 1128;  
 Best Local Similarity 100.0%; Pred. No. 4.3e-14;  
 Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGAAAAGACAGCTATCGGATTCGATGCGACTGGCTGGTTCGCTACCGTGGCCCGAG 60  
 Db 1 ATGAAAAGACAGCTATCGGATTCGATGCGACTGGCTGGTTCGCTACCGTGGCCCGAG 60  
 Qy 61 GCGGCC 66  
 Db 61 GCGGCC 66

RESULT 4  
 AAZ57599  
 ID AAZ57599 standard; DNA; 1574 BP.  
 XX  
 AC AAZ57599;  
 XX  
 DT 19-APR-2000 (first entry)  
 XX  
 DE Erythrocyte and Hepatitis B dual-specific antibody DNA sequence.  
 XX  
 KW Erythrocyte; Hepatitis B; dual-specific antibody; surface antigen; ds.  
 XX  
 OS Unidentified.  
 XX  
 PN CN1232039-A.  
 XX  
 PD 20-OCT-1999.  
 XX  
 XX 02-APR-1999; 99CN-0103517.  
 XX  
 PD 02-APR-1999; 99CN-0103517.  
 XX  
 XX (NAVA-) NAVAL GEN HOSPITAL PLA.  
 XX  
 PI Chen Y, Wang Y;  
 XX  
 XX WPI; 2000-098467/09.  
 XX  
 PT Genetic engineering double specific antibody and its use -  
 PS  
 PS Claim 1; Page 1-2; 6pp; Chinese.  
 XX  
 CC The present invention describes a dual-specific antibody composed of  
 CC the surface antigens to resist against erythrocyte and hepatitis B,

CC which can be used as the test reagent to detect the surface antigen  
 CC of hepatitis B in blood. The dual-specific antibody is prepared by  
 CC recombination technique in gene engineering. The recombination technique  
 CC includes reforming the surface antigen to resist against erythrocyte  
 CC and hepatitis B to become hybrid antibody genes by shortening the  
 CC joining peptide of single-chain antibody, and assembling them in the  
 CC same expression vector. The gene product can be directly extracted from  
 CC the supernatant of bacterial culture liquid. Its advantages are low  
 CC cost, quick detection and simple operation.  
 XX  
 SQ Sequence 1574 BP; 395 A; 390 C; 419 G; 370 T; 0 other;  
 XX

Query Match 100.0%; Score 66; DB 21; Length 1574;  
 Best Local Similarity 100.0%; Pred. No. 4.6e-14;  
 Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGAAAAGACAGCTATCGGATTCGATGCGACTGGCTGGTTCGCTACCGTGGCCCGAG 60  
 Db 792 ATGAAAAGACAGCTATCGGATTCGATGCGACTGGCTGGTTCGCTACCGTGGCCCGAG 851  
 Qy 61 GCGGCC 66  
 Db 852 GCGGCC 857

RESULT 5  
 AAC66993  
 ID AAC66993 standard; DNA; 3450 BP.  
 XX  
 AC AAC66993;  
 XX  
 DT 28-MAR-2001 (first entry)  
 XX  
 DE Vector pCGMT-1b.  
 XX  
 XX Filamentous phage; protein display; pVII; pIX;  
 KW combinatorial antibody library; ds.  
 XX  
 OS Synthetic.  
 XX  
 PN WO200071694-A1.  
 XX  
 PD 30-NOV-2000.  
 XX  
 XX 24-MAY-2000; 2000WO-US14433.  
 XX  
 PF 25-MAY-1999; 99US-0318786.  
 XX  
 PR (SCRI ) SCRIPPS RES INST.  
 XX  
 XX Janda KD, Wirsching P, Lerner RA, Gao C;  
 XX  
 XX WPI; 2001-032030/04.  
 XX  
 DR Novel filamentous phage encapsulating a genome encoding fusion  
 XX polypeptide comprising exogenous polypeptide fused to amino terminus of  
 PT pVII and pIX proteins, for constructing diverse heterodimeric  
 PT polypeptide array -  
 XX  
 XX Example 1; Page 74-75; 90pp; English.  
 PS  
 XX The present invention describes a filamentous phage encapsulating a  
 CC genome encoding a fusion protein. This fusion protein comprises an  
 CC exogenous protein fused to the amino terminus of a filamentous phage pVII  
 CC or pIX protein. This is useful in the design of proteins for medical,  
 CC industrial, environmental and research applications.  
 XX  
 SQ Sequence 3450 BP; 873 A; 799 C; 891 G; 887 T; 0 other;  
 XX

Query Match 100.0%; Score 66; DB 22; Length 3450;  
 Best Local Similarity 100.0%; Pred. No. 5.6e-14;  
 Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;



CC a portion of CDR is replaced with an EPO mimetic, or which has one or  
 CC more of its CDRs fused to an EPO mimetic, is useful for increasing the  
 CC production of red blood cells, where (I) is contacted with haematopoietic  
 CC stem cells or their progenitors. (I) is useful for diagnostics or  
 CC therapeutics, in cell isolation strategies, and for treating patients  
 CC suffering from deficiency in cell populations caused by disease,  
 CC disorders or treatments related to the suppression of haematopoiesis.  
 CC ABQ73288 to ABQ73377 and ABP51669 to ABP51696 represent sequences used  
 CC in the exemplification of the present invention.

XX SQ Sequence 5149 BP; 1279 A; 1244 C; 1372 G; 1254 T; 0 other;

Query Match 100.0%; Score 66; DB 24; Length 5149;  
 Best Local Similarity 100.0%; Pred. No. 6.2e-14;  
 Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGAAAAGACAGCTATCGGATTCGAGTGGCACTGGCTGGCTTTCGCTACCGTGGCCAG 60  
 Db 2611 ATGAAAAGACAGCTATCGGATTCGAGTGGCACTGGCTGGCTTTCGCTACCGTGGCCAG 2670

Qy 61 GCGGCC 66  
 Db 2671 GCGGCC 2676

RESULT 8  
 ABN84077  
 ID ABN84077 standard; DNA; 5149 BP.

AC ABN84077;

DT 23-SEP-2002 (first entry)

DE Phagemid vector pRL8.

KW Vector; phagemid; pRL8; pComb; dimerisation domain; Jun;  
 KW leucine zipper; antibody; gene; ds.

OS Unidentified.

FH Key Location/Qualifiers

FT misc\_feature 2677..3866

FT /\*tag= a

FT /note= "light chain stuffer region"

FT misc\_feature 3968..4337

FT /\*tag= b

FT /note= "heavy chain stuffer region"

FT CDS 4338..4583

FT /\*tag= c

FT /partial

FT /product= "Jun dimerisation domain"

FT /note= "the CDS does not include a start codon"

XX WO200246434-A2.

XX 13-JUN-2002.

XX 07-DEC-2001; 2001WO-US46514.

XX 08-DEC-2000; 2000US-254410P.

XX (ALEX-) ALEXION PHARM INC.

XX Bowdish KS, Barbas-Frederickson S;

XX WPI; 2002-537568/57.

XX P-PSDB; ABB79463.

XX Novel pComb family phagemid comprising a nucleic acid encoding a  
 PT dimerization domain, useful for producing antibody fragments which  
 PT require dimerization in order to crosslink for activation of target  
 PT receptors

XX

PS Disclosure; Fig 8A-C; 17pp; English.

XX The present sequence is of plasmid pRL8, a pComb vector family  
 CC member. pRL8 was constructed by insertion of a PCR fragment  
 CC containing a flexible linker (murine kappa hinge region) followed  
 CC by a Jun leucine zipper dimerisation domain into a pRL4 backbone  
 CC that had been treated to remove the stuffer region. The resulting  
 CC construct was then cut with restriction endonucleases to replace  
 CC scFV with a pRL4 stuffer region for the light and heavy chain. The  
 CC vector is useful in the production of biologically active proteins  
 CC such as antibody fragments which require dimerisation for their  
 CC biological activity. It can also be used in the production and  
 CC screening of libraries made in accordance with phage display  
 CC technology. Homo-dimerisation of single chain antibodies is  
 CC accomplished by the presence of the Jun leucine zipper domain,  
 CC which is responsible for protein-protein interactions. Subcloning  
 CC to pRL8 individually or en masse following FACS sorting or panning  
 CC allows expression, e.g., of dimeric soluble Fabs.

XX SQ Sequence 5149 BP; 1279 A; 1244 C; 1372 G; 1254 T; 0 other;

Query Match 100.0%; Score 66; DB 24; Length 5149;  
 Best Local Similarity 100.0%; Pred. No. 6.2e-14;  
 Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGAAAAGACAGCTATCGGATTCGAGTGGCACTGGCTGGCTTTCGCTACCGTGGCCAG 60  
 Db 2611 ATGAAAAGACAGCTATCGGATTCGAGTGGCACTGGCTGGCTTTCGCTACCGTGGCCAG 2670

Qy 61 GCGGCC 66  
 Db 2671 GCGGCC 2676

RESULT 9

AAL41116

ID AAL41116 standard; DNA; 5683 BP.

XX AC AAL41116;

XX DT 16-OCT-2002 (first entry)

XX DE Plasmid pRL5 nucleic acid sequence.

XX KW Plasmid; vector; replication; expression; foreign genetic; bacteria;  
 KW cyanobacterium; Escherichia coli; pComb3X plasmid; pRL5; pRL5 asc-CAT;  
 KW pRL5-CAT; pRL5 bsi-CAT; gene; ds.

OS Unidentified.

FH Key Location/Qualifiers

FT CDS 591..1451

FT /\*tag= a

FT /product= "Beta lactamase"

FT CDS 2611..2676

FT /\*tag= b

FT /product= "Omp A leader amino acid sequence"

FT CDS 3508..3828

FT /\*tag= c

FT /product= "Kappa constant region amino acid sequence"

FT CDS 3859..3924

FT /\*tag= d

FT /product= "pel B leader amino acid sequence"

FT CDS 4740..5651

FT /\*tag= e

FT /product= "CH1, His6 tag, HA tag, and gene III amino acid  
 sequence"

XX WO200246436-A2.

XX 13-JUN-2002.

XX 07-DEC-2001; 2001WO-US46516.

XX PF

```
XX 08-DEC-2000; 2000US-254411P.
XX (ALEX-) ALEXION PHARM INC.
XX Bowdish KS, Barbas-frederickson S, Wild M, McWhirter J;
XX WPI; 2002-537570/57.
XX P-PSDB; AAO22536, AAO22537, AAO22538, AAO22539, AAO22542.
XX Novel plasmid useful in cloning and expression of foreign genetic
XX information -
XX Claim 1; Fig 4A-B; 39pp; English.
XX The invention relates to a plasmid with a nucleic acid sequence of 6122,
XX 6125, or 5683 base pairs fully defined in the specification. The
XX invention more specifically relates to novel vectors capable of
XX replication and expression of foreign genetic information in bacteria,
XX such as, for example, cyanobacterium and E. coli. The new vectors have
XX been designed to overcome certain drawbacks of the pComb3X plasmid. These
XX new vectors include pRL5, pRL5-CAT, pRL5 asc-CAT and pRL5 bsi-CAT. This
XX polynucleotide sequence represents the plasmid pRL5 nucleic acid
XX of the invention.
XX Sequence 5683 BP; 1337 A; 1495 C; 1540 G; 1311 T; 0 other;
XX
XX Query Match 100.0%; Score 66; DB 24; Length 5683;
XX Best Local Similarity 100.0%; Pred. No. 6.4e-14;
XX Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX Qy 1 ATGAAAAGACAGCTATCGGATTGCGTGCAGTGCAGTGGCTGTTTCGTACCGTGGCCCGAG 60
XX |
XX Db 2611 ATGAAAAGACAGCTATCGGATTGCGTGCAGTGCAGTGGCTGTTTCGTACCGTGGCCCGAG 2670
XX |
XX Qy 61 GCGGCC 66
XX |
XX Db 2671 GCGGCC 2676
XX |
XX
XX RESULT 10
XX AAL41112
XX ID AAL41112 standard; DNA; 6122 BP.
XX AC AAL41112;
XX
XX DT 16-OCT-2002 (first entry)
XX
XX DE Plasmid pRL5 CAT nucleic acid sequence.
XX
XX KW Plasmid; vector; replication; expression; foreign genetic; bacteria;
XX cyanobacterium; Escherichia coli; pComb3X plasmid; pRL5; pRL5 asc-CAT;
XX pRL5-CAT; pRL5 bsi-CAT; gene; ds.
XX
XX OS Unidentified.
XX
XX PH Key Location/Qualifiers
XX CDS 1013..1672
XX FT /*tag= a
XX FT /product= "CAT amino acid sequence"
XX FT 3050..3115
XX CDS /*tag= b
XX FT /product= "Omp A leader amino acid sequence"
XX FT 3953..4267
XX FT /*tag= c
XX FT /product= "Kappa constant region amino acid sequence"
XX FT 4298..4363
XX CDS /*tag= d
XX FT /product= "pel B leader amino acid sequence"
XX FT 5179..6090
XX CDS /*tag= e
XX FT /product= "CHI, His6 tag, HA tag, and gene III amino acid
XX sequence"
```

```
XX WO200246436-A2.
XX
XX PD 13-JUN-2002.
XX
XX 07-DEC-2001; 2001WO-US46516.
XX
XX 08-DEC-2000; 2000US-254411P.
XX (ALEX-) ALEXION PHARM INC.
XX Bowdish KS, Barbas-frederickson S, Wild M, McWhirter J;
XX WPI; 2002-537570/57.
XX P-PSDB; AAO22535, AAO22536, AAO22537, AAO22538, AAO22539.
XX Novel plasmid useful in cloning and expression of foreign genetic
XX information -
XX Claim 1; Fig 7A-E; 39pp; English.
XX The invention relates to a plasmid with a nucleic acid sequence of 6122,
XX 6125, or 5683 base pairs fully defined in the specification. The
XX invention more specifically relates to novel vectors capable of
XX replication and expression of foreign genetic information in bacteria,
XX such as, for example, cyanobacterium and E. coli. The new vectors have
XX been designed to overcome certain drawbacks of the pComb3X plasmid. These
XX new vectors include pRL5, pRL5-CAT, pRL5 asc-CAT and pRL5 bsi-CAT. This
XX polynucleotide sequence represents the plasmid pRL5 nucleic acid
XX of the invention.
XX Sequence 6122 BP; 1449 A; 1571 C; 1639 G; 1463 T; 0 other;
XX
XX Query Match 100.0%; Score 66; DB 24; Length 6122;
XX Best Local Similarity 100.0%; Pred. No. 6.5e-14;
XX Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX Qy 1 ATGAAAAGACAGCTATCGGATTGCGTGCAGTGCAGTGGCTGTTTCGTACCGTGGCCCGAG 60
XX |
XX Db 3050 ATGAAAAGACAGCTATCGGATTGCGTGCAGTGCAGTGGCTGTTTCGTACCGTGGCCCGAG 3109
XX |
XX Qy 61 GCGGCC 66
XX |
XX Db 3110 GCGGCC 3115
XX |
XX
XX RESULT 11
XX AAL41115
XX ID AAL41115 standard; DNA; 6122 BP.
XX AC AAL41115;
XX
XX DT 16-OCT-2002 (first entry)
XX
XX DE Plasmid pRL5 bsi-CAT nucleic acid sequence.
XX
XX KW Plasmid; vector; replication; expression; foreign genetic; bacteria;
XX cyanobacterium; Escherichia coli; pComb3X plasmid; pRL5; pRL5 asc-CAT;
XX pRL5-CAT; pRL5 bsi-CAT; gene; ds.
XX
XX OS Unidentified.
XX
XX PH Key Location/Qualifiers
XX CDS 1013..1672
XX FT /*tag= a
XX FT /product= "CAT amino acid sequence"
XX FT 3050..3115
XX CDS /*tag= b
XX FT /product= "Omp A leader amino acid sequence"
XX FT 3947..4267
XX CDS /*tag= c
XX FT /product= "Kappa constant region amino acid sequence"
XX FT 4298..4363
```



```

FT FT      /*tag= d
FT FT      /product= "pel B leader amino acid sequence"
FT FT      5179..6090
FT FT      /*tag= e
FT FT      /product= "CHI, His6 tag, HA tag, and gene III amino acid
FT FT      sequence"
XX XX      WO200246436-A2.
XX XX      13-JUN-2002.
XX XX      07-DEC-2001; 2001WO-US46516.
XX XX      08-DEC-2000; 2000US-254411P.
XX XX      (ALEX-) ALEXION PHARM INC.
XX XX      Bowdish KS, Barbas-Frederickson S, Wild M, McWhirter J;
XX XX      WPI; 2002-537570/57.
XX XX      P-PSDB; AAO22535, AAO22536, AAO22538, AAO22539, AAO22541.
XX XX      Novel plasmid useful in cloning and expression of foreign genetic
XX XX      information -
XX XX      Disclosure; Fig 11A-E; 39pp; English.
XX XX      The invention relates to a plasmid with a nucleic acid sequence of 6122,
XX XX      6125, or 5683 base pairs fully defined in the specification. The
XX XX      invention more specifically relates to novel vectors capable of
XX XX      replication and expression of foreign genetic information in bacteria,
XX XX      such as, for example, cyanobacterium and E. coli. The new vectors have
XX XX      been designed to overcome certain drawbacks of the pComb3X plasmid. These
XX XX      new vectors include pRL5, pRL5-CAT, pRL5 asc-CAT and pRL5 bsi-CAT. This
XX XX      polynucleotide sequence represents the plasmid pRL5 bsi-CAT nucleic acid
XX XX      sequence of the invention.
XX XX      Query Match 100.0%; Score 66; DB 24; Length 6122;
XX XX      Best Local Similarity 100.0%; Pred. No. 6.5e-14;
XX XX      Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX XX      Qy 1 ATGAAAAGACAGCTATCGGATTCAGTGCATGCTGCTGCTTTCGCTACCGTGGCCGAG 60
XX XX      Db 3050 ATGAAAAGACAGCTATCGGATTCAGTGCATGCTGCTGCTTTCGCTACCGTGGCCGAG 3109
XX XX      Qy 61 GCGGCC 66
XX XX      Db 3110 GCGGCC 3115
XX XX      RESULT 12
XX XX      ABN84078
XX XX      ID ABN84078 standard; DNA; 6122 BP.
XX XX      AC ABN84078;
XX XX      DT 23-SEP-2002 (first entry)
XX XX      DE Plasmid pRL5-CAT.
XX XX      KW Plasmid pRL5-CAT; vector; antibody; gene; ds.
XX XX      OS Unidentified.
XX XX      Key Location/Qualifiers
XX XX      FT CDS 591..800
XX XX      FT FT /*tag= a
XX XX      FT FT /partial
XX XX      FT FT /product= "Amp frag"
XX XX      FT FT /note= "the CDS does not include a stop codon"
XX XX      1013..1672

```

```

FT FT      /*tag= b
FT FT      /product= "Chloramphenicol transferase"
FT FT      complement (2052..2651)
FT FT      /*tag= c
FT FT      /note= "ori"
FT FT      2831..3046
FT FT      /*tag= d
FT FT      /note= "lac promoter"
FT FT      3009..3029
FT FT      /*tag= e
FT FT      /note= "lac rep site"
FT FT      3036..3039
FT FT      /*tag= f
FT FT      3050..3115
FT FT      /*tag= g
FT FT      /partial
FT FT      /product= "OmpA leader"
FT FT      /note= "the CDS does not include a stop codon"
FT FT      3121..3946
FT FT      /*tag= h
FT FT      /note= "light chain variable region stuffer"
FT FT      3951..4269
FT FT      /*tag= i
FT FT      /note= "Kappa Cns"
FT FT      4298..4363
FT FT      /*tag= j
FT FT      /partial
FT FT      /product= "pelB leader"
FT FT      /note= "the CDS does not include a stop codon"
FT FT      4385..5147
FT FT      /*tag= k
FT FT      /note= "heavy chain variable region stuffer"
FT FT      5485..5556
FT FT      /*tag= l
FT FT      /partial
FT FT      /product= "Linker-His tag-HA tag"
FT FT      /note= "the CDS does not include a start codon"
FT FT      5557..6090
FT FT      /*tag= m
FT FT      /partial
FT FT      /gene= "Gene III"
FT FT      /note= "the CDS does not include a start codon"
XX XX      WO200246435-A2.
XX XX      13-JUN-2002.
XX XX      05-DEC-2001; 2001WO-US47452.
XX XX      05-DEC-2000; 2000US-251440P.
XX XX      (ALEX-) ALEXION PHARM INC.
XX XX      Bowdish KS, Barbas-Frederickson S, Lin Y, Renshaw M, Wild M;
XX XX      McWhirter J;
XX XX      WPI; 2002-537569/57.
XX XX      Plasmid for in situ production of genes, comprises two template
XX XX      annealing sequences, such as downstream primer and upstream collar
XX XX      sequence and a restriction site located between the annealing sequences
XX XX      Disclosure; Fig 4A-T; 65pp; English.
XX XX      The present sequence is that of plasmid pRL5-CAT, a derivative of
XX XX      pComb 3X which has been modified to contain chloramphenicol
XX XX      resistance. The invention relates to the engineering of plasmids,
XX XX      of which pRL5-CAT is a preferred example, for in situ production of
XX XX      genes. It was found that nucleic acids encoding a polypeptide
XX XX      can be directly incorporated into a plasmid by DNA polymerisation
XX XX      or by reverse transcription of a nucleic acid template. Preferably,
XX XX      nucleic acids encoding at least a portion of an antibody can be

```

CC directly incorporated into the plasmid by reverse transcription of  
 CC mRNA. The plasmids are engineered to contain 2 template annealing  
 CC sequences (see ABN84079-80 and ABN84082-83), i.e. a downstream  
 CC primer that anneals to a first portion of a nucleic acid template,  
 CC e.g. mRNA encoding at least a portion of an antibody, an upstream  
 CC collar sequence that anneals to a second portion of the template,  
 CC and at least 1 restriction site located between the 2 template  
 CC annealing sequences. A single-stranded DNA plasmid vector is  
 CC produced containing a nucleic acid encoding at least a portion of  
 CC a polypeptide, e.g. a light chain and/or a heavy chain of an  
 CC antibody. This vector can be transformed into a host cell and  
 CC amplified.

SQ Sequence 6122 BP; 1449 A; 1571 C; 1639 G; 1463 T; 0 other;

Query Match 100.0%; Score 66; DB 24; Length 6122;  
 Best Local Similarity 100.0%; Pred. No. 6.5e-14;  
 Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGAAAAGACAGCTATCGGATTCGAGTGCAGTGGCTGGCTTTCGCTACCGTGGCCGAG 60  
 |||||  
 Db 3050 ATGAAAAGACAGCTATCGGATTCGAGTGCAGTGGCTGGCTTTCGCTACCGTGGCCGAG 3109  
 |||||

Qy 61 GCGGCC 66  
 |||||  
 Db 3110 GCGGCC 3115

RESULT 13  
 AAL41114  
 ID AAL41114 standard; DNA; 6125 BP.  
 AC AAL41114;  
 XX  
 DT 16-OCT-2002 (first entry)  
 DE Plasmid pRL5 CAT-Asc nucleic acid sequence.  
 KW Plasmid; vector; replication; expression; foreign genetic; bacteria;  
 KW cyanobacterium; Escherichia coli; pComb3X plasmid; pRL5; pRL5 asc-CAT;  
 KW pRL5-CAT; pRL5 bsi-CAT; gene; ds.  
 XX  
 OS Unidentified.  
 FH Key Location/Qualifiers  
 FT CDS 1013..1672  
 FT /\*tag= a  
 FT /product= "CAT amino acid sequence"  
 FT CDS 3050..3115  
 FT /\*tag= b  
 FT /product= "Omp A leader amino acid sequence"  
 FT CDS 3953..4270  
 FT /\*tag= c  
 FT /product= "Kappa constant region amino acid sequence"  
 FT CDS 4301..4366  
 FT /\*tag= d  
 FT /product= "pel B leader amino acid sequence"  
 FT CDS 5182..6093  
 FT /\*tag= e  
 FT /product= "CHI, His6 tag, HA tag, and gene III amino acid  
 FT sequence"  
 XX WO200246436-A2.  
 XX  
 PD 13-JUN-2002.  
 XX  
 XX 07-DEC-2001; 2001WO-US46516.  
 XX  
 PR 08-DEC-2000; 2000US-254411P.  
 XX  
 PA (ALEX-) ALEXION PHARM INC.  
 XX  
 PI Bowdish KS, Barbas-frederickson S, Wild M, McWhirter J;

XX WPI: 2002-537570/57.  
 DR P-PSDB; AAO22535, AAO22536, AAO22538, AAO22539, AAO22540.  
 XX Novel plasmid useful in cloning and expression of foreign genetic  
 FT information -  
 XX  
 PS Claim 1; Fig 9A-E; 39pp; English.  
 XX  
 CC The invention relates to a plasmid with a nucleic acid sequence of 6122,  
 CC 6125, or 5683 base pairs fully defined in the specification. The  
 CC invention more specifically relates to novel vectors capable of  
 CC replication and expression of foreign genetic information in bacteria,  
 CC such as, for example, cyanobacterium and E. coli. The new vectors have  
 CC been designed to overcome certain drawbacks of the pComb3X plasmid. These  
 CC new vectors include pRL5, pRL5-CAT, pRL5 asc-CAT and pRL5 bsi-CAT. This  
 CC polynucleotide sequence represents the plasmid pRL5 CAT-Asc nucleic acid  
 CC sequence of the invention.

SQ Sequence 6125 BP; 1447 A; 1573 C; 1644 G; 1461 T; 0 other;

Query Match 100.0%; Score 66; DB 24; Length 6125;  
 Best Local Similarity 100.0%; Pred. No. 6.5e-14;  
 Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGAAAAGACAGCTATCGGATTCGAGTGCAGTGGCTGGCTTTCGCTACCGTGGCCGAG 60  
 |||||  
 Db 3050 ATGAAAAGACAGCTATCGGATTCGAGTGCAGTGGCTGGCTTTCGCTACCGTGGCCGAG 3109  
 |||||

Qy 61 GCGGCC 66  
 |||||  
 Db 3110 GCGGCC 3115

RESULT 14  
 AAO92547  
 ID AAO92547 standard; DNA; 6166 BP.  
 AC AAO92547;  
 XX  
 DT 11-MAR-1996 (first entry)  
 DE Expression vector, pPho-TT.  
 XX  
 KW Human; Fab; variable chain; heavy; light; region; VH; VI; HIV; gp120;  
 KW 3b1; 3b3; 3b4; 3b9; MT4; humanised; monoclonal antibody; MAb;  
 KW immunoreaction; neutralisation; passive immunotherapy; tetanus toxin;  
 KW alkaline phosphatase; phoA; ss; cyclic.  
 XX  
 OS Synthetic.  
 XX WO9511317-A1.  
 XX  
 PD 27-APR-1995.  
 XX  
 PF 19-OCT-1994; 94WO-US11907.  
 XX  
 PR 19-SEP-1994; 94US-0308841.  
 PR 19-OCT-1993; 93US-0139409.  
 PR 26-APR-1994; 94US-0233619.  
 XX  
 PA (SCRI ) SCRIPPS RES INST.  
 XX  
 XX Barbas CF, Burton DR, Lerner RA;  
 FI  
 DR WPI; 1995-170235/22.  
 XX  
 XX Synthetic human neutralising monoclonal antibodies to human  
 PT immunodeficiency virus - used for diagnosis and immuno:therapy of  
 FT HIV-induced disease  
 XX  
 XX Example 2; Page 193-197; 249pp; English.  
 PS  
 XX

CC This sequence represents the expression vector, pPho-TT which is a  
 CC modified version of the phagemid expression vector, pComb3 given in  
 CC AAQ92546. pPho-TT provides for the expression of soluble Fabs which are  
 CC secreted into the periplasmic space which is regulated from the alkaline  
 CC phosphatase (phoA) promoter. This plasmid was used within the scope of  
 CC the invention to express various mutagenised human Fab's which comprise  
 CC heavy and light variable regions which bind to HIV gp120. pPho-TT  
 CC consists of a DNA molecule having two cassettes to express two soluble  
 CC proteins a heavy chain and a light chain. The vector comprises,  
 CC operatively linked 5' to 3', a first cassette consisting of the phoA  
 CC promoter/operator sequences, an EcoRI restriction site, a ribosome  
 CC binding site (RBS), an OmpA leader, a SfiI restriction site, a spacer  
 CC region, a cloning region bordered by 5' sacI and 3' XbaI restriction  
 CC sites, an NcoI restriction site between the two cassettes, and a second  
 CC cassette consisting of an expression control RBS, a PelB leader, a human  
 CC consensus amino terminus spacer region comprising the sequence EVQLLE,  
 CC a cloning region bordered by 5' XhoI and 3' SpeI restriction sites  
 CC followed by a SfiI site, expression control stop sequences and a NotI  
 CC restriction site. The pPho-TT expression vector contains a light  
 CC chain stuffer that is 1200 bp in length and a heavy chain stuffer that  
 CC is 300 bp in length. The nucleotide sequences of the heavy and light  
 CC chain stuffers encode the heavy and light chain variable domains of a  
 CC tetanus toxin-specific Fab.

SQ Sequence 6166 BP; 1416 A; 1706 C; 1629 G; 1415 T; 0 other;

Query Match 100.0%; Score 66; DB 16; Length 6166;  
 Best Local Similarity 100.0%; Pred. No. 6.5e-14;  
 Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGAAAAGACAGCTATCGCGATTGCGAGTGCAGTGCACCTGGCTGTTTCGCTACCGTGGCCAG 60  
 Db 4541 ATGAAAAGACAGCTATCGCGATTGCGAGTGCAGTGCACCTGGCTGTTTCGCTACCGTGGCCAG 4600

Qy 61 GCGGCC 66  
 |||||

Db 4601 GCGGCC 4606

#### RESULT 15

AAQ10569

ID AAQ10569 standard; DNA; 73 BP.

AC AAQ10569;

DT 26-APR-1991 (first entry)

DE Omp A signal peptide-encoding sequence.

KW Mirabilis antiviral protein; MAP; Omp A signal sequence; ds.

OS Mirabilis jalapa.

PN EP414134-A.

PD 27-FEB-1991.

PF 16-AUG-1990; 90EP-0115718.

PR 17-AUG-1989; 89JP-0210767.

PA (NISB ) JAPAN TOBACCO INC.

PI Habuka N, Akiyama K, Tsuge H, Mateumoto T, Noma M;

DR WPI; 1991-059386/09.

XX Gene encoding Mirabilis Antiviral Protein and OmpA signal peptide -  
 PT for mass-prodn. of extracellular MAP.

PS Claim 1; page 8; 15pp; English.

CC This Omp A signal peptide-encoding sequence is contained in a

CC recombinant plasmid, flanked on its 3' side by a Mirabilis  
 CC antiviral protein (MAP) gene. The plasmid is used to transform  
 CC E.coli host cells in the large-scale prodn. of MAP. The 5' end  
 CC overhangs the 3' end of the complementary strand by TA and the 5'  
 CC end of the complementary strand overhangs the 3' end of this sense  
 CC strand by CTAG. See also AAQ10568.

SQ Sequence 73 BP; 16 A; 20 C; 21 G; 16 T; 0 other;

Query Match 91.2%; Score 60.2; DB 12; Length 73;

Best Local Similarity 95.4%; Pred. No. 2.7e-12;

Matches 62; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ATGAAAAGACAGCTATCGCGATTGCGAGTGCAGTGCACCTGGCTGTTTCGCTACCGTGGCCAG 60

Db 2 ATGAAAAGACAGCTATCGCGATTGCGAGTGCAGTGCACCTGGCTGTTTCGCTACCGTGGCCAG 61

Qy 61 GCGGC 65

|||

Db 62 GCGGC 66

Search completed: August 18, 2003, 23:28:53

Job time : 24.2513 secs

**THIS PAGE BLANK (USPTO)**

Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	66	100.0	3450	4	US-09-318-786-20	Sequence 20, Appl
2	66	100.0	4891	3	US-08-591-632-43	Sequence 43, Appl
3	66	100.0	4691	3	US-09-611-451-43	Sequence 43, Appl
4	66	100.0	6166	3	US-08-591-632-51	Sequence 51, Appl
5	66	100.0	6166	4	US-09-611-451-51	Sequence 51, Appl
6	60.2	91.2	77	3	US-07-854-845B-2	Sequence 2, Appl
7	60.2	91.2	73	1	US-08-666-354A-10	Sequence 10, Appl
8	60.2	91.2	249	1	US-08-439-132-1	Sequence 1, Appl
9	60.2	91.2	249	6	5223407-1	Patent No. 5223407
10	60.2	91.2	548	3	US-08-828-741B-1	Sequence 1, Appl
11	60.2	91.2	548	4	US-09-160-567-1	Sequence 1, Appl
12	60.2	91.2	548	4	US-09-710-299-1	Sequence 1, Appl
13	59.6	90.3	458	2	US-08-880-829-21	Sequence 21, Appl
14	59.6	90.3	470	3	US-08-828-741B-10	Sequence 10, Appl
15	59.6	90.3	470	4	US-09-160-567-10	Sequence 10, Appl
16	59.6	90.3	470	4	US-09-710-299-10	Sequence 10, Appl
17	59.6	90.3	599	3	US-08-828-741B-7	Sequence 7, Appl
18	59.6	90.3	599	4	US-09-160-567-7	Sequence 7, Appl
19	59.6	90.3	599	4	US-09-710-299-7	Sequence 7, Appl
20	59.6	90.3	1031	3	US-08-828-741B-5	Sequence 5, Appl
21	59.6	90.3	1031	4	US-09-160-567-5	Sequence 5, Appl
22	59.6	90.3	1031	4	US-09-710-299-5	Sequence 5, Appl
23	59.6	90.3	1490	3	US-08-828-741B-3	Sequence 3, Appl
24	59.6	90.3	1490	4	US-09-160-567-3	Sequence 3, Appl
25	59.6	90.3	1490	4	US-09-710-299-3	Sequence 3, Appl
26	59.2	89.7	116	1	US-07-842-089E-23	Sequence 23, Appl
27	59.2	89.7	116	1	US-08-264-48S-23	Sequence 23, Appl

ADDRESSER: The Scripps Research Institute, Office of  
ADDRESSEE: Patent Counsel  
STREET: 10550 No. 6261558th Torrey Pines Road, TPC 8  
CITY: La Jolla  
STATE: CA

COUNTRY: USA

ZIP: 92037

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/591,632

FILING DATE: 19-OCT-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/US94/11907

FILING DATE: 19-OCT-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/308,841

FILING DATE: 19-SEP-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/233,619

FILING DATE: 26-APR-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/139,409

FILING DATE: 19-OCT-1993

ATTORNEY/AGENT INFORMATION:

NAME: Fitting, Thomas

REGISTRATION NUMBER: 34,163

REFERENCE/DOCKET NUMBER: TSRI 332.3

TELEPHONE: (619) 784-2937

TELEFAX: (619) 784-9399

INFORMATION FOR SEQ ID NO: 43:

SEQUENCE CHARACTERISTICS:

LENGTH: 4691 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: circular

MOLECULE TYPE: DNA (genomic)

US-08-591-632-43

Query Match 100.0%; Score 66; DB 3; Length 4691;

Best Local Similarity 100.0%; Pred. No. 1.1e-13;

Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGAAAAGACAGCTATCGGATTGCGACTGCGTGGTTCGCTACCGTGGCCAG 60

Db 2611 ATGAAAAGACAGCTATCGGATTGCGACTGCGTGGTTCGCTACCGTGGCCAG 2670

Qy 61 GCGGCC 66

Db 2671 GCGGCC 2676

### RESULT 3

US-09-611-451-43

Sequence 43, Application US/09611451

Patent No. 6395275

GENERAL INFORMATION:

APPLICANT: Barbas, Carlos F.

Burton, Dennis R.

Lerner, Richard A.

TITLE OF INVENTION: SYNTHETIC HUMAN NEUTRALIZING MONOCLONAL

ANTIBODIES TO HUMAN IMMUNODEFICIENCY VIRUS

NUMBER OF SEQUENCES: 92

CORRESPONDENCE ADDRESS:

ADDRESSEE: The Scripps Research Institute, Office of

Patent Counsel

STREET: 10550 No. 6395275th Torrey Pines Road, TPC 8

CITY: La Jolla

STATE: CA

COUNTRY: USA

ZIP: 92037

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/611,451

FILING DATE: 06-Jul-2000

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/591,632

FILING DATE: 2001-10-29

APPLICATION NUMBER: US 08/308,841

FILING DATE: 19-SEP-1994

APPLICATION NUMBER: US 08/233,619

FILING DATE: 26-APR-1994

APPLICATION NUMBER: US 08/139,409

FILING DATE: 19-OCT-1993

ATTORNEY/AGENT INFORMATION:

NAME: Fitting, Thomas

REGISTRATION NUMBER: 34,163

REFERENCE/DOCKET NUMBER: TSRI 332.3

TELEPHONE: (619) 784-2937

TELEFAX: (619) 784-9399

INFORMATION FOR SEQ ID NO: 43:

SEQUENCE CHARACTERISTICS:

LENGTH: 4691 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: circular

MOLECULE TYPE: DNA (genomic)

SEQUENCE DESCRIPTION: SEQ ID NO: 43:

US-09-611-451-43

Query Match 100.0%; Score 66; DB 4; Length 4691;

Best Local Similarity 100.0%; Pred. No. 1.1e-13;

Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGAAAAGACAGCTATCGGATTGCGACTGCGTGGTTCGCTACCGTGGCCAG 60

Db 2611 ATGAAAAGACAGCTATCGGATTGCGACTGCGTGGTTCGCTACCGTGGCCAG 2670

Qy 61 GCGGCC 66

Db 2671 GCGGCC 2676

### RESULT 4

US-08-591-632-51

Sequence 51, Application US/08591632

Patent No. 6261558

GENERAL INFORMATION:

APPLICANT: Barbas, Carlos F.

Burton, Dennis R.

Lerner, Richard A.

TITLE OF INVENTION: SYNTHETIC HUMAN NEUTRALIZING MONOCLONAL

ANTIBODIES TO HUMAN IMMUNODEFICIENCY VIRUS

NUMBER OF SEQUENCES: 92

CORRESPONDENCE ADDRESS:

ADDRESSEE: The Scripps Research Institute, Office of

Patent Counsel

STREET: 10550 No. 6261558th Torrey Pines Road, TPC 8

CITY: La Jolla

STATE: CA

COUNTRY: USA

ZIP: 92037

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0

;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/591.632  
;; FILING DATE: 19-OCT-1994  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: PCT/US94/11907  
;; FILING DATE: 19-OCT-1994  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 08/308.841  
;; FILING DATE: 19-SEP-1994  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 08/233.619  
;; FILING DATE: 26-APR-1994  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 08/139.409  
;; FILING DATE: 19-OCT-1993  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Fitting, Thomas  
;; REGISTRATION NUMBER: 34,163  
;; REFERENCE/DOCKET NUMBER: TSRI 332.3  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (619) 784-2937  
;; TELEFAX: (619) 784-9399  
;; LENGTH: 6166 base pairs  
;; TYPE: nucleic acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: circular  
;; MOLECULE TYPE: DNA (genomic)  
US-08-591-632-51

Query Match 100.0%; Score 66; DB 3; Length 6166;  
Best Local Similarity 100.0%; Pred. No. 1.2e-13;  
Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGAAAAGACAGCTATCGGATTCGAGTGCAGTGCCTGCTGCTTTCGCTACCGTGGCCAG 60  
Db 4541 ATGAAAAGACAGCTATCGGATTCGAGTGCAGTGCCTGCTGCTTTCGCTACCGTGGCCAG 4600

Qy 61 GCGGCC 66  
Db 4601 GCGGCC 4606

RESULT 5  
US-09-611-451-51  
; Sequence 51, Application US/09611451  
; Patent No. 6395275  
; GENERAL INFORMATION:  
; APPLICANT: Barbas, Carlos F.  
; Burton, Dennis R.  
; Lerner, Richard A.  
; TITLE OF INVENTION: SYNTHETIC HUMAN NEUTRALIZING MONOCLONAL  
; ANTIBODIES TO HUMAN IMMUNODEFICIENCY VIRUS  
; NUMBER OF SEQUENCES: 92  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: The Scripps Research Institute, Office of  
; Patent Counsel  
; STREET: 10550 No. 6395275th Torrey Pines Road, TPC 8  
; CITY: La Jolla  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 92037  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/611.451  
; FILING DATE: 06-Jul-2000  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/591.632

;; FILING DATE: 2001-10-29  
;; APPLICATION NUMBER: US 08/308.841  
;; FILING DATE: 19-SEP-1994  
;; APPLICATION NUMBER: US 08/233.619  
;; FILING DATE: 26-APR-1994  
;; APPLICATION NUMBER: US 08/139.409  
;; FILING DATE: 19-OCT-1993  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Fitting, Thomas  
;; REGISTRATION NUMBER: 34,163  
;; REFERENCE/DOCKET NUMBER: TSRI 332.3  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (619) 784-2937  
;; TELEFAX: (619) 784-9399  
;; INFORMATION FOR SEQ ID NO: 51:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 6166 base pairs  
;; TYPE: nucleic acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: circular  
;; MOLECULE TYPE: DNA (genomic)  
;; SEQUENCE DESCRIPTION: SEQ ID NO: 51:  
US-09-611-451-51

Query Match 100.0%; Score 66; DB 4; Length 6166;  
Best Local Similarity 100.0%; Pred. No. 1.2e-13;  
Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGAAAAGACAGCTATCGGATTCGAGTGCAGTGCCTGCTGCTTTCGCTACCGTGGCCAG 60  
Db 4541 ATGAAAAGACAGCTATCGGATTCGAGTGCAGTGCCTGCTGCTTTCGCTACCGTGGCCAG 4600

Qy 61 GCGGCC 66  
Db 4601 GCGGCC 4606

RESULT 6  
US-07-854-845B-2  
; Sequence 2, Application US/07854845B  
; Patent No. 5340732  
; GENERAL INFORMATION:  
; APPLICANT: HABUKA, No. 5340732iyuki  
; APPLICANT: MIYANO, Masashi  
; APPLICANT: MATSUMOTO, Takashi  
; APPLICANT: NOMA, Masana  
; TITLE OF INVENTION: ANTIVIRAL PROTEIN  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: BIRCH, STEWART, KOLASCH, & BIRCH  
; STREET: 8110 Gatehouse Road, Suite 500 East  
; CITY: Falls Church  
; STATE: VA  
; COUNTRY: USA  
; ZIP: 22042  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/854.845B  
; FILING DATE: 20-MAR-1992  
; CLASSIFICATION: 530  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Murphy Jr., Gerald M  
; REGISTRATION NUMBER: 28,977  
; REFERENCE/DOCKET NUMBER: 42-194P  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 703-205-8000  
; TELEFAX: 703-205-8050  
; TELEX: 248345  
; INFORMATION FOR SEQ ID NO: 2:

## SEQUENCE CHARACTERISTICS:

LENGTH: 73 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
US-07-854-845B-2

Query Match 91.2%; Score 60.2; DB 1; Length 73;

Best Local Similarity 95.4%; Pred. No. 4.9e-12;

Matches 62; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ATGAAAAGACAGCTATCGGATTGCGAGTTCAGTGCGACTGGCTGGTTTCGCTACCGTGCCCGAG 60

Db 2 ATGAAAAGACAGCTATCGGATTGCGAGTTCAGTGCGACTGGCTGGTTTCGCTACCGTGCCCGAG 61

QY 61 GCGGC 65

Db 62 GCGGC 66

## RESULT 7

US-08-666-354A-10

; Sequence 10, Application US/08666354A

; Patent No. 6040141

; GENERAL INFORMATION:

; APPLICANT: KLAUSER, THOMAS

; APPLICANT: KRAMER, JOACHIM

; APPLICANT: MEYER, THOMAS F.

; APPLICANT: POHLNER, JOHANNES

; TITLE OF INVENTION: BACTERIA USED TO PRODUCE STABLE FUSION

; TITLE OF INVENTION: PROTEINS AND METHOD FOR THEIR IDENTIFICATION

; NUMBER OF SEQUENCES: 18

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: BIRCH, STEWART, KOLASCH &amp; BIRCH, LLP.

; STREET: P.O. BOX 747

; CITY: FALLS CHURCH

; STATE: VA

; COUNTRY: USA

; ZIP: 22040-0747

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/666.354A

; FILING DATE: 23-SEP-1996

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: SVENSSON, LEONARD R.

; REGISTRATION NUMBER: 30,330

; REFERENCE/DOCKET NUMBER: 147-157P

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 703-205-8000

; TELEFAX: 703-205-8050

; INFORMATION FOR SEQ ID NO: 10:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 77 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: other nucleic acid

; DESCRIPTION: /desc = "FRAGMENT OF PLASMID pUK165"

US-08-666-354A-10

## Query Match

Best Local Similarity 95.4%; Score 60.2; DB 3; Length 77;

Matches 62; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ATGAAAAGACAGCTATCGGATTGCGAGTTCAGTGCGACTGGCTGGTTTCGCTACCGTGCCCGAG 60

Db 1 ATGAAAAGACAGCTATCGGATTGCGAGTTCAGTGCGACTGGCTGGTTTCGCTACCGTGCCCGAG 60

QY 61 GCGGC 65

Db 61 GCGGC 65

## RESULT 8

US-08-439-132-1

; Sequence 1, Application US/08439132

; Patent No. 5646015

; GENERAL INFORMATION:

; APPLICANT: WONG, W. K. R.

; APPLICANT: SUTHERLAND, MARGARET L.

; TITLE OF INVENTION: EXCRETION OF HETEROLOGOUS PROTEINS FROM

; TITLE OF INVENTION: E. COLI

; NUMBER OF SEQUENCES: 5

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Foley &amp; Lardner

; STREET: Suite 500, 3000 K Street

; CITY: Washington

; STATE: D.C.

; COUNTRY: USA

; ZIP: 20007-5109

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/439,132

; FILING DATE: 11-MAY-1995

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Bent, Stephen A.

; REGISTRATION NUMBER: 29,768

; REFERENCE/DOCKET NUMBER: 16777/200/ALLE

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (202) 672-5300

; TELEFAX: (202) 672-5399

; TELEX: 904136

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 249 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: double

; TOPOLOGY: linear

; US-08-439-132-1

Query Match 91.2%; Score 60.2; DB 1; Length 249;

Best Local Similarity 95.4%; Pred. No. 6.2e-12;

Matches 62; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ATGAAAAGACAGCTATCGGATTGCGAGTTCAGTGCGACTGGCTGGTTTCGCTACCGTGCCCGAG 60

Db 102 ATGAAAAGACAGCTATCGGATTGCGAGTTCAGTGCGACTGGCTGGTTTCGCTACCGTGCCCGAG 161

QY 61 GCGGC 65

Db 162 GCGGC 166

## RESULT 9

5223407-1

; Patent No. 5223407

; APPLICANT: WONG, RAYMOND W.K.; SUTHERLAND, MARGARET L.

; TITLE OF INVENTION: EXCRETION OF HETEROLOGOUS PROTEINS

; FROM E. COLI

; NUMBER OF SEQUENCES: 6

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/07/395,797

; FILING DATE: 18-AUG-1989



```
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 239,145
; FILING DATE: 31-AUG-1988
; SEQ ID NO:1:
; LENGTH: 249
5223407-1

Query Match          91.2%; Score 60.2; DB 6; Length 249;
Best Local Similarity 95.4%; Pred. No. 6.2e-12;
Matches 62; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ATGAAAAGACAGCTATCGGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTAGCCGAG 60
Db 102 ATGAAAAGACAGCTATCGGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTAGCCGAG 60
Qy 61 GCCGC 65
Db 162 GCCGC 166

RESULT 10
US-08-828-741B-1
; Sequence 1, Application US/08828741B
; Patent No. 6043069
; GENERAL INFORMATION:
; APPLICANT: Koentgen, Frank
; APPLICANT: Sues, Gabriele M.
; APPLICANT: Tarlinton, David M.
; APPLICANT: Treutlein, Herbert R.
; TITLE OF INVENTION: CATALYTIC ANTIBODIES AND A METHOD OF
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: United States of America
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/828,741B
; FILING DATE: 26-MAR-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Digiglio, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 10591
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 742-4343
; TELEFAX: (516) 742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 548 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..548
US-08-828-741B-1

Query Match          91.2%; Score 60.2; DB 3; Length 548;
Best Local Similarity 95.4%; Pred. No. 7.2e-12;
Matches 62; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ATGAAAAGACAGCTATCGGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTAGCCGAG 60
Db 1 ATGAAAAGACAGCTATCGGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTAGCCGAG 60
Qy 61 GCCGC 65
Db 162 GCCGC 166

RESULT 11
US-09-160-567-1
; Sequence 1, Application US/09160567
; Patent No. 6326179
; GENERAL INFORMATION:
; APPLICANT: Koentgen, Frank
; APPLICANT: Sues, Gabriele M.
; APPLICANT: Tarlinton, David M.
; APPLICANT: Treutlein, Herbert R.
; TITLE OF INVENTION: CATALYTIC ANTIBODIES AND A METHOD OF
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: United States of America
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/160,567
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/828,741
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Digiglio, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 10591
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 742-4343
; TELEFAX: (516) 742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 548 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..548
US-09-160-567-1

Query Match          91.2%; Score 60.2; DB 4; Length 548;
Best Local Similarity 95.4%; Pred. No. 7.2e-12;
Matches 62; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ATGAAAAGACAGCTATCGGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTAGCCGAG 60
Db 1 ATGAAAAGACAGCTATCGGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTAGCCGAG 60
Qy 61 GCCGC 65
Db 162 GCCGC 166

RESULT 12
```

```

RESULT 13
US-08-880-829-21
; Sequence 21, Application US/08808029
; Patent No. 5925559
; GENERAL INFORMATION:
; APPLICANT: Collins, John
; APPLICANT: Roetgen, Peter
; TITLE OF INVENTION: A Collection of Phagemids, A
; TITLE OF INVENTION: A Collection of Escherichia Coli

```

**RESULT 14**





GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 19, 2003, 04:46:44 ; Search time 18.6243 Seconds  
(without alignments)  
7930.701 Million cell updates/sec

Title: US-09-987-455-6

Perfect score: 66

Sequence: 1 atgaaaagacagctatcgc.....ctaccgtggccaggcgcc 66

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 1504479 seqs, 1118970152 residues

Total number of hits satisfying chosen parameters: 3008958

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:

- 1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq.\*
- 2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq.\*
- 3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq.\*
- 4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq.\*
- 5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq.\*
- 6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq.\*
- 7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq.\*
- 8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq.\*
- 9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq.\*
- 10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq.\*
- 11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq.\*
- 12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq.\*
- 13: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq.\*
- 14: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq.\*
- 15: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq.\*
- 16: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq.\*
- 17: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	66	100.0	66	11	US-09-987-457-1
2	66	100.0	66	11	US-09-987-455-3
3	66	100.0	66	11	US-09-987-455-6
4	66	100.0	1128	11	US-09-987-455-2
5	66	100.0	1128	11	US-09-987-455-5
6	66	100.0	5149	14	US-10-006-593-60
7	66	100.0	6122	14	US-10-006-591-1
8	60.2	91.2	548	12	US-10-345-618-1
9	59.6	90.3	470	12	US-10-345-618-10
10	59.6	90.3	599	12	US-10-345-618-7
11	59.6	90.3	932	9	US-09-809-517A-39
12	59.6	90.3	1031	12	US-10-345-618-5
13	59.6	90.3	1479	12	US-10-345-618-15
14	59.6	90.3	1490	12	US-10-345-618-3
15	59.6	90.3	1574	9	US-09-809-517A-38
16	59.6	90.3	4425	9	US-09-809-517A-40

17	59.2	89.7	4145	14	US-10-001-934-36	Sequence 36, Appl
18	59.2	89.7	5020	14	US-10-001-934-35	Sequence 35, Appl
19	59.2	89.7	5079	9	US-09-809-517A-41	Sequence 41, Appl
20	58.8	89.1	63	9	US-09-760-008A-4	Sequence 4, Appl
21	58.8	89.1	63	14	US-10-076-117-3	Sequence 3, Appl
22	58.8	89.1	63	14	US-10-192-294-4	Sequence 4, Appl
23	58.8	89.1	65	9	US-09-875-494-20	Sequence 20, Appl
24	58.8	89.1	108	11	US-09-848-616-8	Sequence 8, Appl
25	58.8	89.1	256	10	US-09-916-230-13	Sequence 13, Appl
26	58.8	89.1	256	11	US-09-848-616-18	Sequence 18, Appl
27	58.8	89.1	261	10	US-09-916-230-15	Sequence 15, Appl
28	58.8	89.1	261	11	US-09-848-616-20	Sequence 20, Appl
29	58.8	89.1	4614	9	US-09-912-165-17	Sequence 17, Appl
30	58.8	89.1	4657	9	US-09-912-165-18	Sequence 18, Appl
31	58	87.9	585	10	US-09-848-585-28	Sequence 28, Appl
32	58	87.9	585	10	US-09-848-585-30	Sequence 30, Appl
33	58	87.9	981	9	US-09-334-477-36	Sequence 36, Appl
34	58	87.9	990	9	US-09-334-477-38	Sequence 38, Appl
35	58	87.9	2321	10	US-09-995-396-2	Sequence 2, Appl
36	58	87.9	2337	10	US-09-995-396-3	Sequence 3, Appl
37	58	87.9	7083	10	US-09-995-396-1	Sequence 1, Appl
38	57.6	87.3	819	12	US-09-226-157-3	Sequence 3, Appl
39	57.2	86.7	903	14	US-10-033-399B-11	Sequence 11, Appl
40	55.4	83.9	93	13	US-10-004-832-3	Sequence 3, Appl
41	52.8	80.0	867	11	US-09-782-397-16	Sequence 16, Appl
C 42	52.8	80.0	867	11	US-09-782-397-18	Sequence 18, Appl
43	52.8	80.0	918	11	US-09-782-397-13	Sequence 13, Appl
C 44	52.8	80.0	918	11	US-09-782-397-15	Sequence 15, Appl
45	50	75.8	102	14	US-10-006-593-107	Sequence 107, App

#### ALIGNMENTS

#### RESULT 1

US-09-987-457-1  
; Sequence 1, Application US/09987457  
; Publication No. US20030013150A1  
; GENERAL INFORMATION:  
; APPLICANT: Manosroi, Aranya  
; APPLICANT: Manosroi, Aranya  
; APPLICANT: Tayapiwatana, Chatchai  
; APPLICANT: Goetz, Friedrich  
; APPLICANT: Werner, Rolf-Guenther  
; TITLE OF INVENTION: Methods for Large Scale Protein Production in Prokaryotes  
; FILE REFERENCE: 0652.218001  
; CURRENT APPLICATION NUMBER: US/09/987,457  
; CURRENT FILING DATE: 2001-11-14  
; PRIOR APPLICATION NUMBER: 60/268,573  
; PRIOR FILING DATE: 2001-02-15  
; PRIOR APPLICATION NUMBER: GB 00 27 782.2  
; PRIOR FILING DATE: 2000-11-14  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 66  
; TYPE: DNA  
; ORGANISM: Escherichia coli  
US-09-987-457-1

Query Match 100.0%; Score 66; DB 11; Length 66;  
Best Local Similarity 100.0%; Pred. No. 4.6e-16;  
Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGAAAAGACAGCTATCGGATTCGAGTCGACTGCTGCTTTCGTTACGTCGCCAG 60  
|||||  
Db 1 ATGAAAAGACAGCTATCGGATTCGAGTCGACTGCTGCTTTCGTTACGTCGCCAG 60  
|||||

Qy 61 GCGGCC 66  
|||||

Db 61 GCGGCC 66  
|||||

```
RESULT 2
US-09-987-455-3
; Sequence 3, Application US/09987455
; Publication No. US20030049729A1
; GENERAL INFORMATION:
; APPLICANT: Aranya Manosroi
; APPLICANT: Jiradej Manosroi
; APPLICANT: Chatchai Tayapiwatana
; APPLICANT: Friedrich Goetz
; APPLICANT: Rolf-Guenther Werner
; TITLE OF INVENTION: Methods for Large Scale Production of Recombinant
; FILE REFERENCE: 0652.2190001
; CURRENT APPLICATION NUMBER: US/09/987,455
; PRIOR FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: 60/268,574
; PRIOR FILING DATE: 2001-02-15
; PRIOR APPLICATION NUMBER: GB 0027779.8
; PRIOR FILING DATE: 2000-11-14
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 66
; TYPE: DNA
; ORGANISM: Escherichia coli
US-09-987-455-3

Query Match      100.0%; Score 66; DB 11; Length 66;
Best Local Similarity 100.0%; Pred. No. 4.6e-16;
Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 ATGAAAAGACAGCTATCGGATTCGAGTGGCACTGGCTGGTTTCGCTACCGTGCCCCAG 60
Db      1 ATGAAAAGACAGCTATCGGATTCGAGTGGCACTGGCTGGTTTCGCTACCGTGCCCCAG 60
Qy      61 GCGGCC 66
Db      61 GCGGCC 66

RESULT 3
US-09-987-455-6
; Sequence 6, Application US/09987455
; Publication No. US20030049729A1
; GENERAL INFORMATION:
; APPLICANT: Aranya Manosroi
; APPLICANT: Jiradej Manosroi
; APPLICANT: Chatchai Tayapiwatana
; APPLICANT: Friedrich Goetz
; APPLICANT: Rolf-Guenther Werner
; TITLE OF INVENTION: Methods for Large Scale Production of Recombinant
; FILE REFERENCE: 0652.2190001
; CURRENT APPLICATION NUMBER: US/09/987,455
; PRIOR FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: 60/268,574
; PRIOR FILING DATE: 2001-02-15
; PRIOR APPLICATION NUMBER: GB 0027779.8
; PRIOR FILING DATE: 2000-11-14
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 66
; TYPE: DNA
; ORGANISM: Escherichia coli
US-09-987-455-6

Query Match      100.0%; Score 66; DB 11; Length 66;
Best Local Similarity 100.0%; Pred. No. 4.6e-16;
Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 ATGAAAAGACAGCTATCGGATTCGAGTGGCACTGGCTGGTTTCGCTACCGTGCCCCAG 60
Db      1 ATGAAAAGACAGCTATCGGATTCGAGTGGCACTGGCTGGTTTCGCTACCGTGCCCCAG 60
Qy      61 GCGGCC 66
Db      61 GCGGCC 66

RESULT 4
US-09-987-455-2
; Sequence 2, Application US/09987455
; Publication No. US20030049729A1
; GENERAL INFORMATION:
; APPLICANT: Aranya Manosroi
; APPLICANT: Jiradej Manosroi
; APPLICANT: Chatchai Tayapiwatana
; APPLICANT: Friedrich Goetz
; APPLICANT: Rolf-Guenther Werner
; TITLE OF INVENTION: Methods for Large Scale Production of Recombinant
; FILE REFERENCE: 0652.2190001
; CURRENT APPLICATION NUMBER: US/09/987,455
; PRIOR FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: 60/268,574
; PRIOR FILING DATE: 2001-02-15
; PRIOR APPLICATION NUMBER: GB 0027779.8
; PRIOR FILING DATE: 2000-11-14
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 1128
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: coding
; OTHER INFORMATION: sequence for OmpA-K2S fusion protein
US-09-987-455-2

Query Match      100.0%; Score 66; DB 11; Length 1128;
Best Local Similarity 100.0%; Pred. No. 9.7e-16;
Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 ATGAAAAGACAGCTATCGGATTCGAGTGGCACTGGCTGGTTTCGCTACCGTGCCCCAG 60
Db      1 ATGAAAAGACAGCTATCGGATTCGAGTGGCACTGGCTGGTTTCGCTACCGTGCCCCAG 60
Qy      61 GCGGCC 66
Db      61 GCGGCC 66

RESULT 5
US-09-987-455-5
; Sequence 5, Application US/09987455
; Publication No. US20030049729A1
; GENERAL INFORMATION:
; APPLICANT: Aranya Manosroi
; APPLICANT: Jiradej Manosroi
; APPLICANT: Chatchai Tayapiwatana
; APPLICANT: Friedrich Goetz
; APPLICANT: Rolf-Guenther Werner
; TITLE OF INVENTION: Methods for Large Scale Production of Recombinant
; FILE REFERENCE: 0652.2190001
; CURRENT APPLICATION NUMBER: US/09/987,455
; PRIOR FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: 60/268,574
; PRIOR FILING DATE: 2001-02-15
; PRIOR APPLICATION NUMBER: GB 0027779.8
; PRIOR FILING DATE: 2000-11-14
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 1128
```

```
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: coding
; OTHER INFORMATION: sequence for OmpA-K2S fusion protein
US-09-987-455-5

Query Match      100.0%; Score 66; DB 11; Length 1128;
Best Local Similarity 100.0%; Pred. No. 9.7e-16;
Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGAAAAGACAGCTATCGCATTCGAGTGGCAGTGGCTGGTTTCGCTACCGTGGCCCGAG 60
    |||||||
Db 1 ATGAAAAGACAGCTATCGCATTCGAGTGGCAGTGGCTGGTTTCGCTACCGTGGCCCGAG 60
    |||||||

Qy 61 GCGGCC 66
    |||||
Db 61 GCGGCC 66
    |||||

RESULT 6
US-10-006-593-60
; Sequence 60, Application US/10006593
; Publication No. US20030049683A1
; GENERAL INFORMATION:
; APPLICANT: Bowdish, Katherine S.
; APPLICANT: Frederickson, Shana
; APPLICANT: Renshaw, Mark
; TITLE OF INVENTION: RATIONALLY DESIGNED ANTIBODIES
; FILE REFERENCE: 1087-2
; CURRENT APPLICATION NUMBER: US/10/006,593
; CURRENT FILING DATE: 2001-12-05
; PRIOR APPLICATION NUMBER: US 60/251,448
; PRIOR FILING DATE: 2000-12-05
; PRIOR APPLICATION NUMBER: US 60/288,889
; PRIOR FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/294,068
; PRIOR FILING DATE: 2001-05-29
; NUMBER OF SEQ ID NOS: 118
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 60
; LENGTH: 5149
; TYPE: DNA
; ORGANISM: artificial sequence
; FEATURE:
; OTHER INFORMATION: vector
US-10-006-593-60

Query Match      100.0%; Score 66; DB 14; Length 5149;
Best Local Similarity 100.0%; Pred. No. 1.5e-15;
Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGAAAAGACAGCTATCGCATTCGAGTGGCAGTGGCTGGTTTCGCTACCGTGGCCCGAG 60
    |||||||
Db 2611 ATGAAAAGACAGCTATCGCATTCGAGTGGCAGTGGCTGGTTTCGCTACCGTGGCCCGAG 2670
    |||||||

Qy 61 GCGGCC 66
    |||||
Db 2671 GCGGCC 2676
    |||||

RESULT 7
US-10-006-591-1
; Sequence 1, Application US/10006591
; Publication No. US20030049731A1
; GENERAL INFORMATION:
; APPLICANT: Bowdish, Katherine S.
; APPLICANT: Frederickson, Shana
; APPLICANT: Lin, Ying-Chi
; APPLICANT: Renshaw, Mark
; APPLICANT: Wild, Martha
; APPLICANT: McWhitter, John
; TITLE OF INVENTION: ENGINEERED PLASMIDS AND THEIR USE FOR IN SITU PRODUCTION OF GENES
```

```
; FILE REFERENCE: 1087-3
; CURRENT APPLICATION NUMBER: US/10/006,591
; CURRENT FILING DATE: 2001-12-05
; PRIOR APPLICATION NUMBER: 60/251,440
; PRIOR FILING DATE: 2000-12-05
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 6122
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: vector
US-10-006-591-1

Query Match      100.0%; Score 66; DB 14; Length 6122;
Best Local Similarity 100.0%; Pred. No. 1.5e-15;
Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGAAAAGACAGCTATCGCATTCGAGTGGCAGTGGCTGGTTTCGCTACCGTGGCCCGAG 60
    |||||||
Db 3050 ATGAAAAGACAGCTATCGCATTCGAGTGGCAGTGGCTGGTTTCGCTACCGTGGCCCGAG 3109
    |||||||

Qy 61 GCGGCC 66
    |||||
Db 3110 GCGGCC 3115
    |||||

RESULT 8
US-10-345-618-1
; Sequence 1, Application US/10345618
; Publication No. US20030148484A1
; GENERAL INFORMATION:
; APPLICANT: Koentgen, Frank
; APPLICANT: Suess, Gabriele M.
; APPLICANT: Tarlinton, David M.
; APPLICANT: Treutlein, Herbert R.
; TITLE OF INVENTION: CATALYTIC ANTIBODIES AND A METHOD OF PRODUCING SAME
; FILE REFERENCE: 13474
; CURRENT APPLICATION NUMBER: US/10/345,618
; CURRENT FILING DATE: 2003-01-16
; PRIOR APPLICATION NUMBER: US/09/509,031
; PRIOR FILING DATE: 2000-06-09
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 548
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(546)
; OTHER INFORMATION: Description of Artificial Sequence: LHL nucleotide
; OTHER INFORMATION: sequence
US-10-345-618-1

Query Match      91.2%; Score 60.2; DB 12; Length 548;
Best Local Similarity 95.4%; Pred. No. 1.4e-13;
Matches 62; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ATGAAAAGACAGCTATCGCATTCGAGTGGCAGTGGCTGGTTTCGCTACCGTGGCCCGAG 60
    |||||||
Db 1 ATGAAAAGACAGCTATCGCATTCGAGTGGCAGTGGCTGGTTTCGCTACCGTGGCCCGAG 60
    |||||||

Qy 61 GCGGC 65
    |||
Db 61 GCGGC 65
    |||

RESULT 9
US-10-345-618-10
; Sequence 10, Application US/10345618
; Publication No. US20030148484A1
```

```

; GENERAL INFORMATION:
; APPLICANT: Koentgen, Frank
; APPLICANT: Sues, Gabriele M.
; APPLICANT: Tarlinton, David M.
; APPLICANT: Treutlein, Herbert R.
; TITLE OF INVENTION: CATALYTIC ANTIBODIES AND A METHOD OF PRODUCING SAME
; FILE REFERENCE: 13474
; CURRENT APPLICATION NUMBER: US/10/345,618
; CURRENT FILING DATE: 2003-01-16
; PRIOR APPLICATION NUMBER: US/09/509,031
; PRIOR FILING DATE: 2000-06-09
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 470
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(468)
; OTHER INFORMATION: Description of Artificial Sequence:Kappa
; OTHER INFORMATION: nucleotide sequence
US-10-345-618-10

Query Match          90.3%; Score 59.6; DB 12; Length 470;
Best Local Similarity 93.9%; Pred. No. 2.3e-13;
Matches 62; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 ATGAAAAGACAGCTATCGGATTCAGTGGCAGTGGCTGGTTTCGCTACCGTGGCCGAG 60
Db 1 ATGAAAAGACAGCTATCGGATTCAGTGGCAGTGGCTGGTTTCGCTACCGTGGCCGAG 60

Qy 61 GCGGCC 66
Db 61 GCCGAC 66

RESULT 10
US-10-345-618-7
; Sequence 7, Application US/10345618
; Publication No. US20030148484A1
; GENERAL INFORMATION:
; APPLICANT: Koentgen, Frank
; APPLICANT: Sues, Gabriele M.
; APPLICANT: Tarlinton, David M.
; APPLICANT: Treutlein, Herbert R.
; TITLE OF INVENTION: CATALYTIC ANTIBODIES AND A METHOD OF PRODUCING SAME
; FILE REFERENCE: 13474
; CURRENT APPLICATION NUMBER: US/10/345,618
; CURRENT FILING DATE: 2003-01-16
; PRIOR APPLICATION NUMBER: US/09/509,031
; PRIOR FILING DATE: 2000-06-09
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 599
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(597)
; OTHER INFORMATION: Description of Artificial Sequence:HLH.seq
; OTHER INFORMATION: nucleotide sequence
US-10-345-618-7

Query Match          90.3%; Score 59.6; DB 12; Length 599;
Best Local Similarity 93.9%; Pred. No. 2.5e-13;
Matches 62; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 ATGAAAAGACAGCTATCGGATTCAGTGGCAGTGGCTGGTTTCGCTACCGTGGCCGAG 60
Db 1 ATGAAAAGACAGCTATCGGATTCAGTGGCAGTGGCTGGTTTCGCTACCGTGGCCGAG 60
```

```

Qy 61 GCGGCC 66
Db 61 GCCGAC 66

RESULT 11
US-09-809-517A-39
; Sequence 39, Application US/09809517A
; Patent No. US20020034733A1
; GENERAL INFORMATION:
; APPLICANT: Lohning, Corinna
; TITLE OF INVENTION: No. US20020034733A1el methods for displaying (poly)peptides/prot
; TITLE OF INVENTION: particles via disulfide bonds
; FILE REFERENCE: MORPHO/11
; CURRENT APPLICATION NUMBER: US/09/809,517A
; CURRENT FILING DATE: 2001-03-15
; PRIOR APPLICATION NUMBER: EP 99114072.4
; PRIOR FILING DATE: 1999-07-20
; PRIOR APPLICATION NUMBER: EP 00103551.8
; PRIOR FILING DATE: 2000-02-18
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 39
; LENGTH: 932
; TYPE: DNA
; ORGANISM: artificial sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: expression cassette
US-09-809-517A-39

Query Match          90.3%; Score 59.6; DB 9; Length 932;
Best Local Similarity 93.9%; Pred. No. 2.8e-13;
Matches 62; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 ATGAAAAGACAGCTATCGGATTCAGTGGCAGTGGCTGGTTTCGCTACCGTGGCCGAG 60
Db 266 ATGAAAAGACAGCTATCGGATTCAGTGGCAGTGGCTGGTTTCGCTACCGTGGCCGAG 325

Qy 61 GCGGCC 66
Db 326 GCCGAC 331

RESULT 12
US-10-345-618-5
; Sequence 5, Application US/10345618
; Publication No. US20030148484A1
; GENERAL INFORMATION:
; APPLICANT: Koentgen, Frank
; APPLICANT: Sues, Gabriele M.
; APPLICANT: Tarlinton, David M.
; APPLICANT: Treutlein, Herbert R.
; TITLE OF INVENTION: CATALYTIC ANTIBODIES AND A METHOD OF PRODUCING SAME
; FILE REFERENCE: 13474
; CURRENT APPLICATION NUMBER: US/10/345,618
; CURRENT FILING DATE: 2003-01-16
; PRIOR APPLICATION NUMBER: US/09/509,031
; PRIOR FILING DATE: 2000-06-09
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 1031
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1029)
; OTHER INFORMATION: Description of Artificial Sequence:TLHL nucleotide
; OTHER INFORMATION: sequence
US-10-345-618-5

Query Match          90.3%; Score 59.6; DB 12; Length 1031;
Best Local Similarity 93.9%; Pred. No. 2.9e-13;
```



Matches 62; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
Qy 1 ATGAAAAGACAGCTATCGGATTGCGACTGCGCTGGTTCGCTACCGTGGCCCGAG 60  
Db 1 ATGAAAAGACAGCTATCGGATTGCGACTGCGCTGGTTCGCTACCGTGGCCCGAG 60  
Qy 61 GCGGCC 66  
Db 61 GCCGAC 66

RESULT 13  
US-10-345-618-15  
; Sequence 15, Application US/10345618  
; Publication No. US20030148484A1  
; GENERAL INFORMATION:  
; APPLICANT: Koentgen, Frank  
; APPLICANT: Suess, Gabriele M.  
; APPLICANT: Tarlinton, David M.  
; APPLICANT: Treutlein, Herbert R.  
; TITLE OF INVENTION: CATALYTIC ANTIBODIES AND A METHOD OF PRODUCING SAME  
; FILE REFERENCE: 13474  
; CURRENT APPLICATION NUMBER: US/10/345,618  
; PRIOR FILING DATE: 2003-01-16  
; PRIOR APPLICATION NUMBER: US/09/509,031  
; PRIOR FILING DATE: 2000-06-09  
; NUMBER OF SEQ ID NOS: 16  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 15  
; LENGTH: 1479  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence:ccwTLgI  
; OTHER INFORMATION: nucleotide sequence  
US-10-345-618-15

Query Match 90.3%; Score 59.6; DB 12; Length 1479;  
Best Local Similarity 93.9%; Pred. No. 3.2e-13;  
Matches 62; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
Qy 1 ATGAAAAGACAGCTATCGGATTGCGACTGCGCTGGTTCGCTACCGTGGCCCGAG 60  
Db 28 ATGAAAAGACAGCTATCGGATTGCGACTGCGCTGGTTCGCTACCGTGGCCCGAG 87  
Qy 61 GCGGCC 66  
Db 88 GCCGAC 93

RESULT 14  
US-10-345-618-3  
; Sequence 3, Application US/10345618  
; Publication No. US20030148484A1  
; GENERAL INFORMATION:  
; APPLICANT: Koentgen, Frank  
; APPLICANT: Suess, Gabriele M.  
; APPLICANT: Tarlinton, David M.  
; APPLICANT: Treutlein, Herbert R.  
; TITLE OF INVENTION: CATALYTIC ANTIBODIES AND A METHOD OF PRODUCING SAME  
; FILE REFERENCE: 13474  
; CURRENT APPLICATION NUMBER: US/10/345,618  
; PRIOR FILING DATE: 2003-01-16  
; PRIOR APPLICATION NUMBER: US/09/509,031  
; PRIOR FILING DATE: 2000-06-09  
; NUMBER OF SEQ ID NOS: 16  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 3  
; LENGTH: 1490  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; NAME/KEY: CDS

; LOCATION: (1)...(1488)  
; OTHER INFORMATION: Description of Artificial Sequence:CATAB-TEV  
; OTHER INFORMATION: nucleotide sequence  
US-10-345-618-3

Query Match 90.3%; Score 59.6; DB 12; Length 1490;  
Best Local Similarity 93.9%; Pred. No. 3.2e-13;  
Matches 62; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
Qy 1 ATGAAAAGACAGCTATCGGATTGCGACTGCGCTGGTTCGCTACCGTGGCCCGAG 60  
Db 1 ATGAAAAGACAGCTATCGGATTGCGACTGCGCTGGTTCGCTACCGTGGCCCGAG 60  
Qy 61 GCGGCC 66  
Db 61 GCCGAC 66

RESULT 15  
US-09-809-517A-38  
; Sequence 38, Application US/09809517A  
; Patent No. US20020034733A1  
; GENERAL INFORMATION:  
; APPLICANT: Lohning, Corinna  
; TITLE OF INVENTION: No. US20020034733A1 methods for displaying (poly)peptides/prote  
; FILE REFERENCE: MORPHO/11  
; CURRENT APPLICATION NUMBER: US/09/809,517A  
; CURRENT FILING DATE: 2001-03-15  
; PRIOR APPLICATION NUMBER: EP 99114072.4  
; PRIOR FILING DATE: 1999-07-20  
; PRIOR APPLICATION NUMBER: EP 00103551.8  
; PRIOR FILING DATE: 2000-02-18  
; NUMBER OF SEQ ID NOS: 41  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 38  
; LENGTH: 1574  
; TYPE: DNA  
; ORGANISM: artificial sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: expression cassette  
US-09-809-517A-38

Query Match 90.3%; Score 59.6; DB 9; Length 1574;  
Best Local Similarity 93.9%; Pred. No. 3.2e-13;  
Matches 62; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
Qy 1 ATGAAAAGACAGCTATCGGATTGCGACTGCGCTGGTTCGCTACCGTGGCCCGAG 60  
Db 266 ATGAAAAGACAGCTATCGGATTGCGACTGCGCTGGTTCGCTACCGTGGCCCGAG 325  
Qy 61 GCGGCC 66  
Db 326 GCCGAC 331

Search completed: August 19, 2003, 14:22:28  
Job time : 19.6243 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 18, 2003, 22:44:33 ; Search time 193.082 Seconds  
(without alignments)  
8307.845 Million cell updates/sec

Title: US-09-987-455-6  
Perfect score: 66  
Sequence: 1 atgaaaagacagctatgc.....ctaccgtggccagcgccg 66

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues  
Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST:  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hic:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hic:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: em\_gss\_hum:\*  
18: em\_gss\_inv:\*  
19: em\_gss\_pln:\*  
20: em\_gss\_vrt:\*  
21: em\_gss\_fun:\*  
22: em\_gss\_mam:\*  
23: em\_gss\_mus:\*  
24: em\_gss\_pro:\*  
25: em\_gss\_rod:\*  
26: em\_gss\_phg:\*  
27: em\_gss\_vrl:\*  
28: gb\_gss1:\*  
29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	48	72.7	655	28	AF075794
C 2	27.8	42.1	386	14	CB806580
C 3	26.8	40.6	310	9	AI635506
C 4	26.8	40.6	312	9	AW004057

C 5	26.8	40.6	321	10	BE503913
C 6	26.8	40.6	325	9	AA398159
C 7	26.8	40.6	340	9	AA401699
C 8	26.8	40.6	409	13	BY641012
C 9	26.8	40.6	569	13	BU783334
C 10	26.8	40.6	667	13	EX099539
C 11	26.8	40.6	694	29	CNS04PHI
C 12	26.8	40.6	786	12	BI560274
C 13	26.8	40.6	871	13	BUR853590
C 14	26.6	40.3	268	28	AQ905715
C 15	26.6	40.3	359	9	AU244728
C 16	26.6	40.3	377	9	AU244560
C 17	26.6	40.3	450	10	BG234261
C 18	26.6	40.3	472	28	AQ904535
C 19	26.6	40.3	612	12	BJ031326
C 20	26.6	40.3	645	13	BQ400448
C 21	26.6	40.3	753	14	CA808042
C 22	26.6	40.3	833	14	CB349272
C 23	26.6	40.3	885	13	BQ737041
C 24	26.6	40.3	888	13	BQ737333
C 25	26.6	40.3	889	14	CD101043
C 26	26.6	40.3	897	13	BQ733411
C 27	26.6	40.3	980	14	CB206009
C 28	26.6	40.3	1124	14	CB206058
C 29	26.2	39.7	271	9	AU184837
C 30	26.2	39.7	374	13	BQ094120
C 31	26.2	39.7	695	14	CD355744
C 32	26.2	39.7	973	13	BQ938965
C 33	26	39.4	495	28	AQ833969
C 34	26	39.4	853	29	CNS03TQE
C 35	26	39.4	950	29	CNS04IET
C 36	26	39.4	1016	29	CNS02VHS
C 37	26	39.4	1387	29	BZ557231
C 38	25.6	38.8	291	14	H33495
C 39	25.6	38.8	331	14	CA377775
C 40	25.6	38.8	438	14	CA373964
C 41	25.6	38.8	621	9	AL966295
C 42	25.6	38.8	652	9	AL644617
C 43	25.6	38.8	673	9	AL655729
C 44	25.6	38.8	687	14	CA369617
C 45	25.6	38.8	691	9	AL630991

ALIGNMENTS

RESULT 1  
AF075794/c  
LOCUS AF075794 Salmonella typhimurium LT2, Lambda DASH II Salmonella typhimurium genomic clone 107-T3, genomic survey sequence. 655 bp DNA linear GSS 29-AUG-2000  
DEFINITION AF075794 Salmonella typhimurium LT2, Lambda DASH II Salmonella typhimurium genomic clone 107-T3, genomic survey sequence.  
ACCESSION AF075794  
VERSION AF075794.1 GI:3320664  
KEYWORDS GSS.  
SOURCE Salmonella typhimurium  
ORGANISM Salmonella typhimurium  
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales; Enterobacteriaceae; Salmonella.  
REFERENCE 1 (bases 1 to 655)  
Wong,R.M.-Y., Wong,K.K., Benson,N.R. and McClelland,M.  
TITLE Sample sequencing of a Salmonella typhimurium LT2 lambda library: comparison to the Escherichia coli K12 genome  
JOURNAL FEMS Microbiol. Lett. 173 (2), 411-423 (1999)  
MEDLINE 9243757  
PUBMED 10227170  
COMMENT Contact: McClelland M  
Molecular Biology  
Sidney Kimmel Cancer Center  
3099 Science Park Road, San Diego, CA 92121, USA  
Email: mclelland@lifsci.sdsu.edu  
Class: shotgun.  
Location/Qualifiers  
source 1. .655

```

/organism="Salmonella typhimurium"
/mol_type="genomic DNA"
/strain="LT2"
/db_xref="taxon:602"
/clone="107-T3"
/clone_lib="Salmonella typhimurium LT2, Lambda DASH II"
/note="Vector: Lambda DASH II; sequenced using Li-Cor
sequencer"
BASE COUNT 195 a 158 c 139 g 161 t . 2 others
ORIGIN
Query Match 72.7%; Score 48; DB 28; Length 655;
Best Local Similarity 85.2%; Pred. No. 8.9e-06;
Matches 52; Conservative 1; Mismatches 8; Indels 0; Gaps 0;
QY 5 AAAAGACAGCTATCCGATTGCGATGGCACTGGCTGGTTTCGCTACCGTGGCCAGGCGG 64
Db 298 AAAAGACGACGATTCGATGGCACTGGCTGGTTTCGCTACCGTGGCCAGGCGG 239
QY 65 C 65
Db 238 C 238

```

```

RESULT 2
CB806580/c
LOCUS
DEFINITION
AMGNNUC:SRP2-00054-G9-A srp2 (10238) Rattus norvegicus cDNA clone
srp2-00054-g9 5', mRNA sequence.
ACCESSION
CB806580
VERSION
CB806580.1 GI:29924779
KEYWORDS
EST.
SOURCE
Rattus norvegicus (Norway rat)
ORGANISM
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

```

```

REFERENCE
1 (bases 1 to 386)
Angen EST Program.
AUTHORS
Angen Rat EST Program
TITLE
Unpublished
JOURNAL
Contact: Dan Fitzpatrick
COMMENT
Amgen, Inc
One Amgen Center Drive, Thousand Oaks, CA 91320-1799, USA
Tel: 805 447-4881
Plate: 00054 row: g column: 9.

```

```

FEATURES
source
1..386
Location/Qualifiers
/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"
/clone="srp2-00054-g9"
/tissue_type="peneal gland brain"
/clone_lib="srp2 (10238)"
/note="Vector: pSPOR1; Site_1: Sal1; Site_2: NotI; peneal
gland brain region"
BASE COUNT 57 a 149 c 119 g 60 t 1 others
ORIGIN
Query Match 42.1%; Score 27.8; DB 14; Length 386;
Best Local Similarity 64.1%; Pred. No. 48;
Matches 41; Conservative 0; Mismatches 23; Indels 0; Gaps 0;
QY 2 TGAAGACAGCTATCCGATTGCGATGGCACTGGCTGGTTTCGCTACCGTGGCCAGG 61
Db 85 TGAAGACGACGCGCGCGCGGCGGCGGCGGCGGCGCTAGGNCAGCGACGCGGCGCGG 26
QY 62 CGCG 65
Db 25 CGCG 22

```

```

RESULT 3

```

```

AI635506/c
LOCUS
DEFINITION
ts95a03.x1 NCI_CGAP_GC6 Homo sapiens cDNA clone IMAGE:2238988 3',
mRNA sequence.
ACCESSION
AI635506
VERSION
AI635506.1 GI:4686836
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 310)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
AUTHORS
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL
Unpublished
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgapsb@mail.nih.gov
Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael
R. Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldo, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 383 Std Error: 0.00
Seq primer: -40UP from Gibco.
Location/Qualifiers
1..310
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2238988"
/tissue_type="pooled germ cell tumors"
/lab_host="DH10B"
/clone_lib="NCI_CGAP GC6"
/note="Vector: p773D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; Plasmid DNA
from the normalized library NCI_CGAP GC4 was prepared, and
ss circles were made in vitro. Following HAP purification,
this DNA was used as tracer in a subtractive hybridization
reaction. The driver was PCR-amplified cDNAs from a pool
of 5,000 clones made from the same library (cloneIDs
1257096-1258631, 1469084-1470983, and 1475592-1476743).
Subtraction by Bento Soares and M. Fatima Bonaldo. "
BASE COUNT 71 a 72 c 64 g 103 t
ORIGIN
Query Match 40.6%; Score 26.8; DB 9; Length 310;
Best Local Similarity 64.5%; Pred. No. 97;
Matches 40; Conservative 0; Mismatches 22; Indels 0; Gaps 0;
QY 5 AAAAGACAGCTATCCGATTGCGATGGCACTGGCTGGTTTCGCTACCGTGGCCAGGCGG 64
Db 268 AAAAGACGACATCACCGGCCATTACGCTTGTGCTTTTGTAGCAGAGCCAGCAG 209
QY 65 CC 66
Db 208 CC 207

```

```

RESULT 4
AW004057/c
LOCUS
DEFINITION
wg85e09.x1 NCI_CGAP_GC6 Homo sapiens cDNA clone IMAGE:2478956 3',
mRNA sequence.
ACCESSION
AW004057
VERSION
AW004057.1 GI:5850973
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens

```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

## REFERENCE

1 (bases 1 to 312)

## AUTHORS

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

## TITLE

Tumor Gene Index

## JOURNAL

Unpublished

## COMMENT

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael

R. Emmert-Buck, M.D., Ph.D.

cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima

Bonaldo, Ph.D.

cDNA Library Arrayed by: Greg Lennon, Ph.D.

DNA Sequencing by: Washington University Genome Sequencing Center

Clone distribution: NCI-CGAP clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

www-bio.llnl.gov/bbrp/image/image.html

Seq primer: -40UP from Gibco.

Location/Qualifiers

1..312

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:247885e"

/tissue\_type="pooled germ cell tumors"

/lab\_host="DH10B"

/clone\_lib="NCI CGAP GC6"

/note="Vector: pTTT3D-Pac (Pharmacia) with a modified

polylinker; Site 1: Not I; Site 2: Eco RI; Plasmid DNA

from the normalized library NCI-CGAP GC4 was prepared, and

ss circles were made in vitro. Following HAP purification,

this DNA was used as tracer in a subtractive hybridization

reaction. The driver was PCR-amplified cDNAs from a pool

of 5,000 clones made from the same library (cloneIDs

1257096-1258631, 1469064-1470983, and 1475592-1476743).

Subtraction by Bento Soares and M. Fatima Bonaldo.

72 a 72 c 64 g 104 t

BASE COUNT

ORIGIN

Query Match 40.6%; Score 26.8; DB 9; Length 312;

Best Local Similarity 64.5%; Pred. No. 97;

Matches 40; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

Qy 5 AAAAGACAGTATCGGATTCAGTGGCAGTGGCTGGTTGCTACCTGCGCCAGCGG 64

Db 269 AAAAAGACGATCATCAGGCGCATTCAGCTTTGCTTTGTAGCAGAGCCAGGCG 210

Qy 65 CC 66

Db 209 CC 208

RESULT 5

BE503913/c

LOCUS

DEFINITION

h235e09.x1 NCI CGAP GC6 Homo sapiens cDNA clone IMAGE:3209992 3',

mRNA sequence.

ACCESSION

BE503913

VERSION

BE503913.1 GI:97063221

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 321)

NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index

Unpublished

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael

100

R. Emmert-Buck, M.D., Ph.D.

cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima

Bonaldo, Ph.D.

cDNA Library Arrayed by: Greg Lennon, Ph.D.

DNA Sequencing by: Washington University Genome Sequencing Center

Clone distribution: NCI-CGAP clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL, send email to:

info@image.llnl.gov

Seq primer: -40UP from Gibco

High quality sequence stop: 313.

Location/Qualifiers

1..321

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:3209992"

/tissue\_type="pooled germ cell tumors"

/lab\_host="DH10B"

/clone\_lib="NCI CGAP GC6"

/note="Vector: pTTT3D-Pac (Pharmacia) with a modified

polylinker; Site 1: Not I; Site 2: Eco RI; Plasmid DNA

from the normalized library NCI-CGAP GC4 was prepared, and

ss circles were made in vitro. Following HAP purification,

this DNA was used as tracer in a subtractive hybridization

reaction. The driver was PCR-amplified cDNAs from a pool

of 5,000 clones made from the same library (cloneIDs

1257096-1258631, 1469064-1470983, and 1475592-1476743).

Subtraction by Bento Soares and M. Fatima Bonaldo.

75 a 77 c 65 g 104 t

BASE COUNT

ORIGIN

Query Match 40.6%; Score 26.8; DB 10; Length 321;

Best Local Similarity 64.5%; Pred. No. 98;

Matches 40; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

Qy 5 AAAAGACAGTATCGGATTCAGTGGCAGTGGCTGGTTGCTACCTGCGCCAGCGG 64

Db 269 AAAAAGACGATCATCAGGCGCATTCAGCTTTGCTTTGTAGCAGAGCCAGGCG 210

Qy 65 CC 66

Db 209 CC 208

RESULT 6

AA398159/c

LOCUS

DEFINITION

z60e09.s1 Soares\_testis\_NHT Homo sapiens cDNA clone IMAGE:726760

3', mRNA sequence.

ACCESSION

AA398159

VERSION

AA398159.1 GI:2051405

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 325)

Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S.,

Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin, J., Moore, B.,

Schellenberg, K., Steptoe, M., Tan, F., Theising, B., White, Y., Wyllie

, T., Waterston, R., and Wilson, R.

WashU-Merck EST Project 1997

Unpublished

Contact: Wilson RK

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@wustl.wustl.edu

This clone is available royalty-free through LLNL; contact the

IMAGE Consortium (info@image.llnl.gov) for further information.

Insert Length: 392 Std Error: 0.00

Seq primer: -41m13 fwd. ET from Amersham.

FEATURES  
source

Location/Qualifiers  
1. 325  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="GDB:5923670"  
/db\_xref="taxon:9606"  
/clone="IMAGE:726760"  
/sex="male"  
/lab\_host="DH10B"  
/clone\_lib="Soares testis NHT"  
/note="Vector: pT73D-Pac (Pharmacia) with a modified  
polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA  
was prepared from mRNA obtained from Clontech Laboratories  
, Inc., and primed with a Not I - oligo(dT) primer [5',  
TGTTACCAATCTGAAGTGGAGCGCGCCCAATTTTTTTTTTTT 3'].  
Double-stranded cDNA was ligated to Eco RI adaptors  
(Pharmacia), digested with Not I and cloned into the Not I  
and Eco RI sites of the modified pT73 vector. Library  
went through one round of normalization to Cot5, and was  
constructed by Bento Soares and M. Fatima Bonaldo."

72 a 72 c 64 g 117 t

BASE COUNT  
ORIGIN

Query Match 40.6%; Score 26.8; DB 9; Length 325;  
Best Local Similarity 64.5%; Pred. No. 98;  
Matches 40; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

Qy 5 AAAAGACAGCTATCGGATTGCGACGTGGCTGTTTCTACCTGCCCGCGCGG 64  
Db 282 AAAAAGACGATCATCCAGGCGATTCACGCTTTGTGCTTTGTAGCAGAGCCCGCGAG 223

Qy 65 CC 66  
Db 222 CC 221

## RESULT 7

AA401699  
LOCUS  
DEFINITION  
5', mRNA sequence.  
AA401699.1 GI:2057290  
EST.  
Homo sapiens (human)  
Homo sapiens  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisler, G., Jost, S.,  
Kucaba, F., Lacy, W., Le, N., Lennon, G., Marra, M., Martin, J., Moore, B.,  
Schellenberg, K., Steptoe, M., Tan, F., Theising, B., White, Y., Wylie,  
T., Waterston, R., and Wilson, R.  
WashU-Merck EST Project 1997  
Unpublished  
Contact: Wilson RK  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@watson.wustl.edu  
This clone is available royalty-free through LNL; contact the  
IMAGE Consortium (info@image.llnl.gov) for further information.  
Insert Length: 392 Std Error: 0.00  
Seq primer: -28ml3 rev2 ET from Amersham.

Location/Qualifiers

1. 340  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="GDB:5923670"  
/db\_xref="taxon:9606"  
/clone="IMAGE:726760"  
/sex="male"

## FEATURES

## source

Location/Qualifiers  
1. 325  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="GDB:5923670"  
/db\_xref="taxon:9606"  
/clone="IMAGE:726760"  
/sex="male"

Query Match 40.6%; Score 26.8; DB 9; Length 340;  
Best Local Similarity 64.5%; Pred. No. 1e+02;  
Matches 40; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

Qy 5 AAAAGACAGCTATCGGATTGCGACGTGGCTGTTTCTACCTGCCCGCGCGG 64  
Db 76 AAAAAGACGATCATCCAGGCGATTCACGCTTTGTGCTTTGTAGCAGAGCCCGCGAG 135

Qy 65 CC 66  
Db 136 CC 137

RESULT 8  
BY641012/c  
LOCUS  
DEFINITION  
409 bp mRNA linear EST 15-DEC-2002  
CDNA clone K430341N20 3', mRNA sequence.

ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS

BY641012  
Mus musculus (house mouse)  
Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 409)  
Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S.,  
Nikaido, I., Osato, N., Saito, R., Saito, H., Yamana, A., Schonbach, C.,  
Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Hume, D. A.,  
Gojobori, T., Baldarelli, R., Hill, D. P., Bult, C., Hume, D. A.,  
Quackenbush, J., Schriml, L. M., Kanapin, A., Matsuda, H., Batalov, S.,  
Beisel, K. W., Blake, J. A., Bradt, D., Brusic, V., Chothia, C., Corbani,  
L. E., Cousins, S., Dalla, E., Dragani, T. A., Fletcher, C. F., Forrest,  
A., Frazer, K. S., Gaasterland, T., Gariboldi, M., Gissi, C., Godzik, A.,  
Gough, J., Grimmond, S., Gustincich, S., Hirokawa, N., Jackson, I. J.,  
Jarvis, E. D., Kanai, A., Kawaji, H., Kawasawa, Y., Kedzierski, R. M.,  
King, B. L., Konegaya, A., Kurochkin, I. V., Lee, Y., Lenhard, B., Lyons,  
P. A., Maglott, D. R., Maltais, L., Marchionni, L., McKenzie, L., Miki,  
H., Nagashima, T., Numata, K., Okido, T., Pavan, W. J., Perle, G.,  
Pesole, G., Petrovsky, N., Pillai, R., Pontius, J. U., Qi, D.,  
Ramachandran, S., Ravasi, T., Reed, J. C., Reed, D. J., Reid, J., Ring,  
B. Z., Ringwald, M., Sandelin, A., Schneider, C., Semple, C. A., Setou,  
M., Shimada, K., Sultana, R., Takenaka, Y., Taylor, M. S., Teasdale,  
R. D., Tomita, M., Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y.,  
Watanabe, Y., Wells, C., Wilming, L. G., Wynshaw-Boris, A., Yanagisawa,  
M., Yang, I., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A.,  
Carninci, P., Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura,  
M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K.,  
Arakawa, T., Fukuda, S., Hara, A., Hashizume, M., Imotani, K., Ishii,  
Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata,  
K., Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander,  
E. S., Rogers, J., Birney, E. and Hayashizaki, Y.

Analysis of the mouse transcriptome based on functional annotation  
of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)

JOURNAL  
MEDLINE  
PUBMED  
12466851

CONTACT: Yoshihide Hayashizaki

/lab\_host="DH10B"

/clone\_lib="Soares testis NHT"  
/note="Vector: pT73D-Pac (Pharmacia) with a modified  
polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA  
was prepared from mRNA obtained from Clontech Laboratories  
, Inc., and primed with a Not I - oligo(dT) primer [5',  
TGTTACCAATCTGAAGTGGAGCGCGCCCAATTTTTTTTTTTT 3'].  
Double-stranded cDNA was ligated to Eco RI adaptors  
(Pharmacia), digested with Not I and cloned into the Not I  
and Eco RI sites of the modified pT73 vector. Library  
went through one round of normalization to Cot5, and was  
constructed by Bento Soares and M. Fatima Bonaldo."

106 a 75 c 82 g 77 t

BASE COUNT  
ORIGIN

Query Match 40.6%; Score 26.8; DB 9; Length 340;  
Best Local Similarity 64.5%; Pred. No. 1e+02;  
Matches 40; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

Qy 5 AAAAGACAGCTATCGGATTGCGACGTGGCTGTTTCTACCTGCCCGCGCGG 64  
Db 76 AAAAAGACGATCATCCAGGCGATTCACGCTTTGTGCTTTGTAGCAGAGCCCGCGAG 135

Qy 65 CC 66  
Db 136 CC 137

## RESULT 8

BY641012/c  
LOCUS  
DEFINITION  
409 bp mRNA linear EST 15-DEC-2002  
CDNA clone K430341N20 3', mRNA sequence.

ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS

BY641012  
Mus musculus (house mouse)  
Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 409)  
Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S.,  
Nikaido, I., Osato, N., Saito, R., Saito, H., Yamana, A., Schonbach, C.,  
Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Hume, D. A.,  
Gojobori, T., Baldarelli, R., Hill, D. P., Bult, C., Hume, D. A.,  
Quackenbush, J., Schriml, L. M., Kanapin, A., Matsuda, H., Batalov, S.,  
Beisel, K. W., Blake, J. A., Bradt, D., Brusic, V., Chothia, C., Corbani,  
L. E., Cousins, S., Dalla, E., Dragani, T. A., Fletcher, C. F., Forrest,  
A., Frazer, K. S., Gaasterland, T., Gariboldi, M., Gissi, C., Godzik, A.,  
Gough, J., Grimmond, S., Gustincich, S., Hirokawa, N., Jackson, I. J.,  
Jarvis, E. D., Kanai, A., Kawaji, H., Kawasawa, Y., Kedzierski, R. M.,  
King, B. L., Konegaya, A., Kurochkin, I. V., Lee, Y., Lenhard, B., Lyons,  
P. A., Maglott, D. R., Maltais, L., Marchionni, L., McKenzie, L., Miki,  
H., Nagashima, T., Numata, K., Okido, T., Pavan, W. J., Perle, G.,  
Pesole, G., Petrovsky, N., Pillai, R., Pontius, J. U., Qi, D.,  
Ramachandran, S., Ravasi, T., Reed, J. C., Reed, D. J., Reid, J., Ring,  
B. Z., Ringwald, M., Sandelin, A., Schneider, C., Semple, C. A., Setou,  
M., Shimada, K., Sultana, R., Takenaka, Y., Taylor, M. S., Teasdale,  
R. D., Tomita, M., Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y.,  
Watanabe, Y., Wells, C., Wilming, L. G., Wynshaw-Boris, A., Yanagisawa,  
M., Yang, I., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A.,  
Carninci, P., Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura,  
M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K.,  
Arakawa, T., Fukuda, S., Hara, A., Hashizume, M., Imotani, K., Ishii,  
Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata,  
K., Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander,  
E. S., Rogers, J., Birney, E. and Hayashizaki, Y.

## TITLE

Analysis of the mouse transcriptome based on functional annotation  
of 60,770 full-length cDNAs

## JOURNAL

22354683

## MEDLINE

12466851

## COMMENT

Contact: Yoshihide Hayashizaki

Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute  
The Institute of Physical and Chemical Research (RIKEN)  
1-7-22 Suhiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
Tel: 81-45-503-9222  
Fax: 81-45-503-9216  
Email: genome-res@gsc.riken.go.jp/  
URL: http://genome.gsc.riken.go.jp/  
Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hirozane, T., Imotani, K., Ishii, Y., Itoh, M., Kawai, J., Konno, H., Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Numazaki, R., Ohno, M., Sakai, K., Sakazume, N., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M., Waki, K., Watabiki, A., Muramatsu, M. and Hayashizaki, Y. Direct Submission  
Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)  
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)  
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)  
Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)  
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.  
Tissues were provided by Michela Fagiolini and Takao K. Hensch (Laboratory for Neuronal Circuit Development Brain Science Institute RIKEN 2-1 Hirosawa, Wako-shi, Saitama 351-0198 Japan) whose assistance we gratefully acknowledge.  
Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

## FEATURES

source

```

1. .409
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="K430341N20"
/tissue_type="visual cortex"
/clone_lib="RIKEN full-length enriched, visual cortex"
89 a 96 c 87 g 135 t 2 others

```

BASE COUNT  
ORIGIN

Query Match 40.6%; Score 26.8; DB 13; Length 409;  
Best Local Similarity 64.5%; Pred. No. 1.1e+02;  
Matches 40; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

```

Qy 1 ATGAAAGACAGCTATCGGATTCGACTGGCTGCTGCTTTCGCTACCGTGGCCAG 60
Db 102 AGGAAAGAGAGGCAAGCAGGCTTAGAGAGTGGTGGCTGCTTTCAGCTTCCACAG 43

Qy 61 GC 62
Db 42 GC 41

```

RESULT 9  
BU783334  
LOCUS

DEFINITION BU783334 569 bp mRNA linear EST 11-OCT-2002  
in02g05.y1 Human insulinoma Homo sapiens cDNA clone IMAGE:6123441  
5', mRNA sequence.

ACCESSION BU783334  
VERSION BU783334.1 GI:23827413

KEYWORDS EST.  
SOURCE Homo sapiens (human)

## ORGANISM

Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

## REFERENCE

## AUTHORS

1 (bases 1 to 569)  
Melton, D., Brown, J., Kenty, G., Permutt, A., Lee, C., Kaestner, K.,  
Lemishka, I., Scarce, M., Brestelli, J., Gradwohl, G., Clifton, S.,  
Hallier, L., Marra, M., Pape, D., Wylie, T., Martin, J., Blistain, A.,  
Schmitt, A., Theising, B., Ritter, E., Ronko, I., Bennett, J., Cardenas,  
M., Gibbons, M., McCann, R., Cole, R., Tsagarishvili, R., Williams, T.,  
Jackson, Y. and Bowers, Y.  
Endocrine Pancreas Consortium  
Unpublished  
Other ESTs: in02g05.x1  
Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue  
Endocrine Pancreas Consortium  
Harvard University, Howard Hughes Medical Institute  
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,  
MA 02138  
Tel: 617-495-1812  
Fax: 617-495-8557  
Email: dmelton@bioh.harvard.edu  
Library was constructed by Dr. J. Ferrer In vivo mass-excised to  
pBluescript SK- by Dr. H. Inoue DNA sequencing by: Washington  
University Genome Sequencing Center For information on obtaining a  
clone please contact: Dr. Hiroshi Inoue (hinoue@im.wustl.edu)  
Seq primer: -40RP from Gibco  
High quality sequence stop: 429.

TITLE  
JOURNAL  
COMMENTFEATURES  
source

```

1. .569
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6123441"
/tissue_type="insulinoma"
/lab_host="DHI0B (phage-resistant)"
/clone_lib="Human insulinoma"
/notes="Organ: pancreas; Vector: pBluescript SK-; Site 1:  
XhoI; Site 2: EcoRI; Constructed with lambda ZAPII system  
(Stratagene) by Dr. J. Ferrer, in vivo mass-excised to  
pBluescript SK- by Dr. H. Inoue following the Washington  
University protocol  
(http://genome.wustl.edu/est/lambda\_protocol.shtml).  
Please contact Hiroshi Inoue, MD/PhD for further  
information on this library (Metabolism Division, Permutt  
Laboratory, Washington University School of Medicine, Box  
8127, 660 S Euclid Ave, St. Louis, MO 63110). Note: this  
is a Washington University Pancreas EST project library."
118 a 160 c 144 g 147 t

```

BASE COUNT  
ORIGIN

Query Match 40.6%; Score 26.8; DB 13; Length 569;  
Best Local Similarity 64.5%; Pred. No. 1.2e+02;  
Matches 40; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

```

Qy 5 AAAAGACAGCTATCGGATTCGACTGGCTGCTGCTTTCGCTACCGTGGCCAGCGG 64
Db 253 AAAAAGACAGCCATCACAGGCCATTCACGCTTTGCTTTTGTAGCAGAGCCAGCAG 312

Qy 65 CC 66
Db 313 CC 314

```

RESULT 10  
BX099539  
LOCUS

DEFINITION BX099539 667 bp mRNA linear EST 06-FEB-2003  
in02g05.y1 Soares testis NHT Homo sapiens cDNA clone IMAGE:9986231928  
5', mRNA sequence.

ACCESSION BX099539  
VERSION BX099539.1 GI:27844114

KEYWORDS EST.  
SOURCE Homo sapiens (human)

## ORGANISM

Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 667)

**AUTHORS**  
 Ebert, L., Heil, O., Hennig, S., Neubert, P., Partsch, E., Peters, M.,  
 Radelof, U., Schneider, D., and Korn, B.  
**TITLE**  
 Human Unigeneset - RZPD3  
**JOURNAL**  
 Unpublished  
**COMMENT**  
 Contact: Ina Rofls  
 RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH  
 Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany  
 RZPD: IMAGp998G231928.  
 RZPDLIB: I.M.A.G.E. CDNA Clone Collection;  
 Human Unigeneset - RZPD3 (RZPDLIB No.972)  
 http://www.rzpd.de/CloneCards/cgi-  
 bin/showLib.pl.cgi/response?libNo=972 Contact: Ina Rofls  
 RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH  
 Heubnerweg 6, D-14059 Berlin, Germany  
 Tel.: +49 30 32639 101  
 Fax: +49 30 32639 111  
 www.rzpd.de  
 This clone is available royalty-free from RZPD;  
 contact RZPD (clone@rzpd.de) for further information. Seq primer:  
 M13r, Primer sequence: TTTCACACAGGAACAGCTATGAC.

**FEATURES**  
 source  
 1..667  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGp998G231928 ; IMAGE:782638"  
 /sex="male"  
 /lab\_host="DH10B"  
 /clone\_lib="Soares testis NHT"  
 /note="Vector: pTT3D-Pac (Pharmacia) with a modified  
 polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA  
 was prepared from mRNA obtained from Clontech Laboratories  
 Inc., and primed with a Not I - Oligo(dT) primer [5'  
 TGTTACCAATCTGAAGTGGAGCGGCCCAATTTTTTTTTTTT 3'].  
 Double-stranded cDNA was ligated to Eco RI adaptors  
 (Pharmacia), digested with Not I and cloned into the Not I  
 and Eco RI sites of the modified pT7t3 vector. Library  
 went through one round of normalization to Cot5, and was  
 constructed by Bento Soares and M. Fatima Bonaldo."

**BASE COUNT**  
 195 a 142 c 167 g 162 t 1 others

**Query Match**  
 Best Local Similarity 40.6%; Score 26.8; DB 13; Length 667;  
 Matches 40; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

**Qy**  
 5 AAAGACAGCTATCGGATTCGATGCGACATGCGTGGTTCCTACCGTCCCGGCGG 64  
 |||||  
**Db**  
 43 AAAAAGACGCTATCAGGCGCATTCAGCTTTGTGCTTTGTAGCAGCGCCGCGAG 102  
 |||||

**Qy**  
 65 CC 66  
 ||  
**Db**  
 103 CC 104

**RESULT 11**  
 CNS04PHI/c  
**LOCUS**  
 DEFINITION  
 Tetraodon nigroviridis genome survey sequence PUC-ori end of clone  
 127D11 of library G from Tetraodon nigroviridis, genomic survey  
 sequence.  
 AL301311  
 AL301311.1 GI:8178354  
 GSS; genome survey sequence.  
**KEYWORDS**  
 Tetraodon nigroviridis  
**SOURCE**  
 Tetraodon nigroviridis  
**ORGANISM**  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
 Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;  
 Tetraodontidae; Tetraodontidae; Tetraodon.  
 1

**REFERENCE**  
**AUTHORS**  
 Roest Crolius, H., Jaillon, O., Dasilva, C., Bouneau, L., Fisher, C.,  
 Bernot, A., Fizames, C., Wincker, P., Brottier, P., Quetier, F.,

**TITLE**  
 Saurin, W. and Weissenbach, J.  
 Estimating of human gene number provided by genome-wide analysis  
 using Tetraodon nigroviridis DNA sequence  
 Nat. Genet. 25 (2), 235-238 (2000)  
 20296633  
 10835645  
**REFERENCE**  
**AUTHORS**  
 Roest Crolius, H., Jaillon, O., Dasilva, C., Ozouf-Costaz, C.,  
 Fizames, C., Fischer, C., Bouneau, L., Billault, A., Quetier, F.,  
 Saurin, W., Bernot, A. and Weissenbach, J.  
 Characterization and repeat analysis of the compact genome of the  
 freshwater pufferfish Tetraodon nigroviridis  
 Genome Res. 10 (7), 939-949 (2000)  
 20359837  
 10899143  
**REFERENCE**  
 3 (bases 1 to 694)  
 Genoscope.  
 Direct Submission  
 Submitted (12-APR-2000) Genoscope - Centre National de Sequencage :  
 BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)  
 - Web : www.genoscope.cns.fr)  
 This sequence is a single read and was generated as part of a large  
 scale clone-end sequencing project of the Tetraodon nigroviridis  
 genome. For more information, please take a look at  
 http://www.genoscope.cns.fr/Tetraodon.

**FEATURES**  
 source  
 1..694  
 /organism="Tetraodon nigroviridis"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:99883"  
 /clone="127D11"  
 /clone\_lib="G"  
 /note="Genoscope sequence ID : COBG127CE06SP1-end :  
 PUC-ori"

**BASE COUNT**  
 189 a 122 c 186 g 175 t 22 others

**Query Match**  
 40.6%; Score 26.8; DB 29; Length 694;  
 Best Local Similarity 60.6%; Pred. No. 1.3e+02;  
 Matches 40; Conservative 2; Mismatches 24; Indels 0; Gaps 0;

**Qy**  
 1 ATGAAAGACAGCTATCGGATTCGATGCGACATGCGTGGTTCCTACCGTCCCGCAG 60  
 |||||  
**Db**  
 278 ACCTGATGCGGCTTTGCGAAGCTCTKACATACAKAGACGATGCTGCGGCCAG 219  
 |||||

**Qy**  
 61 GCGGCC 66  
 |||||  
**Db**  
 218 GCGTCC 213

**RESULT 12**  
 BIS0274  
**LOCUS**  
 DEFINITION  
 603253511F1 NIH\_MGC\_97 Homo sapiens cDNA clone IMAGE:5295746 5',  
 mRNA sequence.  
 BIS0274  
 BIS0274.1 GI:15447588  
**VERSION**  
 EST.  
**KEYWORDS**  
 Homo sapiens (human)  
**SOURCE**  
 Homo sapiens  
**ORGANISM**  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 786)  
 NIH-MGC http://mgc.nci.nih.gov/.  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished  
**REFERENCE**  
**AUTHORS**  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgapbs@mail.nih.gov  
 Tissue Procurement: Miklos Palkovits, M.D., Ph.D.  
 cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki  
 Toshiyuki and Piero Carninci (RIKEN)  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.



Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
 Plate: L1AM11748 row: f column: 03  
 High quality sequence stop: 772.  
 Location/Qualifiers

#### FEATURES

1..786  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5295746"  
 /lab\_host="DH10B"  
 /clone\_lib="NIH\_MGC\_97"

/note="Organ: testis; Vector: pBluescriptR (modified pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI (gtcggg); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTVN-3', size-selected for average insert size 2.2 kb and normalized to ROT 5. This is a primary library enriched for full-length clones and constructed using the Cap-trapper method (Carninci, in preparation). Library constructed by M. Brownstein (NIMH/NHGRI, National Institutes of Health). Note: this is a NIH\_MGC Library."

238 a 168 c 187 g 193 t

#### BASE COUNT

##### ORIGIN

Query Match 40.6%; Score 26.8; DB 12; Length 786;  
 Best Local Similarity 64.5%; Pred. No. 1.3e+02;  
 Matches 40; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

Qy 5 A A A A G A C G C T A T C G C G A T T G C A G T G G C A C T G G T T G C T A C C G T G C C C C A G G C G 64

Db 56 A A A A A A C A G C C A T C A C C A G G C C A T T C A C G C T T G T G C T T T T G T T A G C A G A G C C C A G G C A G 115

Qy 65 CC 66

Db 116 CC 117

#### RESULT 13

##### LOCUS

BU853590 871 bp mRNA linear EST 16-OCT-2002  
 AGENCOURT\_10418173 NIH\_MGC\_82 Homo sapiens cDNA clone IMAGE:6620518  
 5', mRNA sequence.

Qy 5 A A A A G A C G C T A T C G C G A T T G C A G T G G C A C T G G T T G C T A C C G T G C C C C A G G C G 64

Db 56 A A A A A A C A G C C A T C A C C A G G C C A T T C A C G C T T G T G C T T T T G T T A G C A G A G C C C A G G C A G 115

Qy 65 CC 66

Db 116 CC 117

#### REFERENCE

##### AUTHORS

##### TITLE

##### JOURNAL

##### COMMENT

Contact: Robert Strausberg, Ph.D.  
 Email: [cgabbs-remail.nih.gov](mailto:cgabbs-remail.nih.gov)  
 Tissue Procurement: CLONTECH  
 cDNA Library Preparation: CLONTECH Laboratories, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Agencourt Biotechnology Corporation  
 Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
 Plate: L1AM11748 row: 1 column: 22  
 High quality sequence stop: 454.  
 Location/Qualifiers

#### FEATURES

##### source

1..871  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:6620518"  
 /lab\_host="DH10B (T1 phage-resistant)"  
 /clone\_lib="NIH\_MGC\_82"

/note="Organ: testis; Vector: pDNR-LIB (Clontech); Site 1: SfiI (ggcgccctcgcc); Site 2: SfiI (ggccattatggcc); 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CACGGCCATTATGGCC-3' and 3' adaptor sequence: 5'-ATTCTAGAGCGGAGCGGCACATG-dt(30)BN-3' (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.35 kb (range 0.9-4.0 kb). 14/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA)."

BASE COUNT 240 a 226 c 203 g 202 t

##### ORIGIN

Query Match 40.6%; Score 26.8; DB 13; Length 871;

Best Local Similarity 64.5%; Pred. No. 1.4e+02;

Matches 40; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

Qy 5 A A A A G A C G C T A T C G C G A T T G C A G T G G C A C T G G T T G C T A C C G T G C C C C A G G C G 64

Db 38 A A A A A A C A G C C A T C A C C A G G C C A T T C A C G C T T T G T T A G C A G A G C C C A G G C A G 97

Qy 65 CC 66

Db 98 CC 99

#### RESULT 14

##### LOCUS

AQ905715/c 268 bp DNA linear GSS 09-JAN-2001  
 DEFINITION GSTc07088 Trypanosoma cruzi random genomic library Trypanosoma cruzi genomic clone G55P8, genomic survey sequence.

Qy 5 A A A A G A C G C T A T C G C G A T T G C A G T G G C A C T G G T T G C T A C C G T G C C C C A G G C G 64

Db 38 A A A A A A C A G C C A T C A C C A G G C C A T T C A C G C T T T G T T A G C A G A G C C C A G G C A G 97

Qy 65 CC 66

Db 98 CC 99

#### REFERENCE

##### AUTHORS

##### TITLE

##### JOURNAL

##### MEDLINE

##### PUBMED

##### COMMENT

On Sep 14, 2000 this sequence version replaced gi:9373324.  
 Contact: Sanchez D.O.  
 Instituto de Investigaciones Biotecnologicas (Univ. Nac. de Gral San Martin)  
 Av. Gral Paz S/N, INTI, Edificio 24, B 1650 KNA, San Martin, Buenos Aires, Argentina  
 Tel: (54-11) 4580/7255/7  
 Fax: (54-11) 4752-9639  
 Email: [deanchez@ib.unsam.edu.ar](mailto:deanchez@ib.unsam.edu.ar)  
 Sequences were basecalled with phred and vector was masked with crossmatch (see <http://genome.washington.edu>). Sequences were then trimmed from both ends to remove low quality bases and masked vector.  
 Seq primer: T7  
 Class: Shotgun.  
 Location/Qualifiers

#### FEATURES

##### source

1..268  
 /organism="Trypanosoma cruzi"  
 /mol\_type="genomic DNA"  
 /strain="CL-Brener"  
 /db\_xref="taxon:5693"  
 /clone="G55P8"  
 /cell\_type="epimastigote"  
 /clone\_lib="Trypanosoma cruzi random genomic library"  
 /note="Vector: pBS(-) (Stratagene); T. cruzi DNA was randomly sheared using a nebulizer and the 1 to 2 kb range was gel purified and cloned into the dephosphorylated

BASE COUNT 60 a 63 c 84 g 61 t  
ORIGIN  
HincII site of the vector"

Query Match 40.3%; Score 26.6; DB 28; Length 268;  
Best Local Similarity 71.4%; Pred. No. 1.1e+02;  
Matches 35; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 16 ATCGCGATTGCAGTGGCGACTGGCTGGTTTCCTACCGTGGCCCGCGG 64  
|||||  
Db 189 ATCGTGTGCTGTGTCGCGGCGCTTTGGCTCCCGTCGCCCTGGCGG 141  
|||||

RESULT 15  
AU244728  
LOCUS AU244728 359 bp mRNA linear EST 21-FEB-2002  
DEFINITION AU244728 Shibata Xenopus AEM lambda-ZAP II cDNA library Xenopus  
laevis cDNA clone p5el2 5', mRNA sequence.  
ACCESSION AU244728  
VERSION AU244728.1 GI:18850654  
KEYWORDS EST.  
SOURCE Xenopus laevis (African clawed frog)  
ORGANISM Xenopus laevis  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;  
Xenopodinae; Xenopus.  
REFERENCE 1 (bases 1 to 359)  
AUTHORS Shibata,M., Itoh,M., Ohmori,S., Shinga,J. and Taira,M.  
TITLE Systematic screening and expression analysis of the head organizer  
genes in Xenopus embryos  
JOURNAL Dev. Biol. 239 (2), 241-256 (2001)  
MEDLINE 21643879  
PubMed 11784032  
COMMENT Contact: Masanori Taira  
Department of Biological Sciences  
Graduate School of Science, University of Tokyo; CREST, Japan  
Science and Technology Corporation, Japan  
7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan  
Tel: 81-03-5841-4434  
Fax: 81-03-5841-4434  
Email: m\_taira@biol.s.u-tokyo.ac.jp,  
URL: http://www.biol.s.u-tokyo.ac.jp/users/lmb/lmb-hp.html.  
Location/Qualifiers  
1. 359  
/organism="Xenopus laevis"  
/mol\_type="mRNA"  
/db\_xref="taxon:8355"  
/clone="p5el2"  
/tissue type="anterior endomesoderm"  
/dev stages="late gastrula to early neurula stages 12.5-13"  
/clone\_lib="Shibata Xenopus AEM lambda-ZAP II cDNA  
library"

BASE COUNT 90 a 92 c 82 g 93 t 2 others  
ORIGIN

Query Match 40.3%; Score 26.6; DB 9; Length 359;  
Best Local Similarity 63.1%; Pred. No. 1.2e+02;  
Matches 41; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY 2 TGAAGAGACAGCTATCGAGTGGCAGTGGCAGTGGCTGGTTTCCTACCGTGGCCCGG 61  
|||||  
Db 129 TGAAGAGGAGATATCGTATCGTATGGGATCGCTGGTTCTCATGCTCTCCAGCAGG 188  
|||||

QY 62 CGGCC 66  
|||||  
Db 189 CAGCC 193

Search completed: August 19, 2003, 08:29:19  
Job time : 194.082 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 8, 2003, 17:08:17 ; Search time 44 Seconds  
(without alignments)  
1359.998 Million cell updates/sec

Title: US-09-987-455-8

Perfect score: 2061

Sequence: 1 MKKTAIAIALAGFAATVAQ.....GVYTKVTNYLDWIRDNRPG 377

Scoring table:

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_19Jun03;\*

1: /SIDSI1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*  
2: /SIDSI1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*  
3: /SIDSI1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*  
4: /SIDSI1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*  
5: /SIDSI1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*  
6: /SIDSI1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*  
7: /SIDSI1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*  
8: /SIDSI1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*  
9: /SIDSI1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*  
10: /SIDSI1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*  
11: /SIDSI1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*  
12: /SIDSI1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*  
13: /SIDSI1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*  
14: /SIDSI1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*  
15: /SIDSI1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*  
16: /SIDSI1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*  
17: /SIDSI1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*  
18: /SIDSI1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*  
19: /SIDSI1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*  
20: /SIDSI1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*  
21: /SIDSI1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*  
22: /SIDSI1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*  
23: /SIDSI1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*  
24: /SIDSI1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2061	100.0	377	23	AAE25034
2	1961	95.1	372	20	AAW87770
3	1961	95.1	396	13	AAAP25190
4	1961	95.1	438	10	AAAP94409
5	1961	95.1	439	16	AAAG68851
6	1961	95.1	446	11	AAAR08150
7	1961	95.1	472	10	AAAP94416
8	1961	95.1	519	16	AAAR68852
9	1961	95.1	527	11	AAAR05489

10	1961	95.1	527	11	AAAR05488	tPA024 precursor p
11	1961	95.1	527	12	AAAR13910	T-PA deriv. (I).
12	1961	95.1	527	12	AAAR13911	T-PA deriv. (II).
13	1961	95.1	527	12	AAAR13912	T-PA deriv. (III).
14	1961	95.1	527	12	AAAR13913	T-PA deriv. (IV).
15	1961	95.1	527	12	AAAR13914	T-PA deriv. (V).
16	1961	95.1	527	13	AAAR20223	t-PA analogue expr
17	1961	95.1	527	13	AAAR21593	tPA variant - E94A
18	1961	95.1	527	13	AAAR21594	tPA variant - D95A
19	1961	95.1	527	13	AAAR21595	tPA variant - D95G
20	1961	95.1	527	13	AAAR21596	tPA variant - E94A
21	1961	95.1	527	13	AAAR21600	tPA variant - E94A
22	1961	95.1	527	13	AAAR22617	Mutated recombinan
23	1961	95.1	527	13	AAAR22618	Mutated recombinan
24	1961	95.1	527	13	AAAR22619	Mutated recombinan
25	1961	95.1	527	13	AAAR22620	Mutated recombinan
26	1961	95.1	527	13	AAAR22621	Mutated recombinan
27	1961	95.1	527	19	AAW45907	Single chain form
28	1961	95.1	527	23	AAE25044	Human tissue plasm
29	1961	95.1	527	23	AAE24190	Human tissue plasm
30	1961	95.1	527	23	AAAG79362	Human tissue plasm
31	1961	95.1	528	16	AAAR68853	Bifibronection doma
32	1961	95.1	531	7	AAAP60700	Sequence of tissue
33	1961	95.1	537	12	AAAR11661	Tissue plasminogen
34	1961	95.1	539	12	AAAR12846	T-PA growth domain
35	1961	95.1	552	12	AAAR11662	Tissue plasminogen
36	1961	95.1	556	12	AAAR13148	T-PA variant contg
37	1961	95.1	557	12	AAAR13020	T-PA variant contg
38	1961	95.1	557	12	AAAR13149	T-PA variant contg
39	1961	95.1	558	12	AAAR13153	T-PA with -ve char
40	1961	95.1	559	12	AAAR12340	T-PA variant contg
41	1961	95.1	559	12	AAAR13151	T-PA with -ve char
42	1961	95.1	560	12	AAAR12341	T-PA variant contg
43	1961	95.1	561	12	AAAR12367	T-PA with -ve char
44	1961	95.1	562	7	AAAP60790	Sequence of human
45	1961	95.1	562	7	AAAP60214	Sequence of active

#### ALIGNMENTS

RESULT 1  
AAE25034  
ID AAE25034 standard; Protein; 377 AA.

XX AC AAE25034;

XX AC AAE25034;

DT 30-OCT-2002 (first entry)

XX OmpA-K2S fusion protein.

DE Tissue plasminogen activator; tPA; kringle 2 serine protease; stroke;

KW K2S; cardiac infarction; acute myocardial infarction; pulmonary embolism;  
KW artery occlusion; deep vein thrombosis; blood clotting; thrombolytic;  
KW cerebroprotective; cardiant; ompA; fusion protein.

OS Unidentified.

XX WO200240650-A2.

XX 23-MAY-2002.

PD 07-NOV-2001; 2001WO-EPI2857.

XX 14-NOV-2000; 2000GB-0027779.

XX (BOEH ) BOEHRINGER INGELHEIM INT GMBH.

PA Werner R, Goetz F, Tayapiwatana C, Manosroi J, Manosroi A;

PI WPI; 2002-519376/55.

XX N-PSDB; AAD40614.

DR

XX

PT Producing active, correctly folded recombinant tissue plasminogen  
 PT activator, Kringle 2 serine protease in prokaryotic cells by expressing  
 PT the protein-encoding DNA operably linked to DNA coding for signal  
 PT peptide.OmpA

PS Claim 25; Page 35-36; 80pp; English.

XX The present invention relates to a method of producing extracellularly  
 CC secreted, active, correctly folded, recombinant tissue plasminogen  
 CC activator (tPA), Kringle 2 serine protease molecule (K2S) or their  
 CC variants in prokaryotic cells by expressing the protein-encoding DNA  
 CC operably linked to DNA coding for signal peptide OmpA. The method is  
 CC useful for producing recombinant DNA-derived tPA, K2S or their variants.  
 CC Sequences of the invention are useful for manufacturing a medicament  
 CC for treating stroke, cardiac infarction, acute myocardial infarction,  
 CC pulmonary embolism, any artery occlusion such as intracranial artery  
 CC occlusion (e.g., arteries supplying the brain), peripherally occluded  
 CC arteries, coronary artery occlusion, deep vein thrombosis or related  
 CC diseases associated with unwanted blood clotting. The present sequence  
 CC is a fusion protein comprising OmpA and K2S protein.

XX Sequence 377 AA;

Query Match 100.0%; Score 2061; DB 23; Length 377;  
 Best Local Similarity 100.0%; Pred. No. 1.2e-167;  
 Matches 377; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKKTAIAIAVALAGFATVAQAASEGNSDCYFGNGSAVRGTHSLTESGASCLPWNMSMLIG 60  
 DB 1 MKKTAIAIAVALAGFATVAQAASEGNSDCYFGNGSAVRGTHSLTESGASCLPWNMSMLIG 60  
 QY 61 KVTYAQNPSAQAALGLGKHNCRNPDGAKPWCHVKNRRLTWECYDVPSCSTCGLROYSQ 120  
 DB 61 KVTYAQNPSAQAALGLGKHNCRNPDGAKPWCHVKNRRLTWECYDVPSCSTCGLROYSQ 120  
 QY 121 POFRIKGLFADIASHPWQAIAFAKHRRSPGERFLCGGILISSCWILSAHCFQERFPPH 180  
 DB 121 POFRIKGLFADIASHPWQAIAFAKHRRSPGERFLCGGILISSCWILSAHCFQERFPPH 180  
 QY 181 HLTVLGRTYRVVPEEEOKEFEVKYIVHKEFDDTDYNDIALQLKSDSRCAQESSVV 240  
 DB 181 HLTVLGRTYRVVPEEEOKEFEVKYIVHKEFDDTDYNDIALQLKSDSRCAQESSVV 240  
 QY 241 RTVCLPPADLQLPDWTECELSYGKHEALSPFYSERLKEAHLVLPSSRCTSOHLNRTV 300  
 DB 241 RTVCLPPADLQLPDWTECELSYGKHEALSPFYSERLKEAHLVLPSSRCTSOHLNRTV 300  
 QY 301 TDNMLCAGDTRSGGPQANLHDACQDGGPLVCLNDGRMTLVGIIISWGLGCGQKDVPGVY 360  
 DB 301 TDNMLCAGDTRSGGPQANLHDACQDGGPLVCLNDGRMTLVGIIISWGLGCGQKDVPGVY 360  
 QY 361 TKVTNYLDWIRDNRPG 377  
 DB 361 TKVTNYLDWIRDNRPG 377

RESULT 2

AAW87770 standard; Protein; 372 AA.

XX AAW87770;

XX 29-MAR-1999 (first entry)

XX Human tissue plasminogen activator.

XX Tissue plasminogen activator-like protease; t-PALP; human;  
 KW circulatory system-related disorder; blood clotting; stroke;  
 KW thrombosis; peripheral arterial occlusion; pulmonary embolism;  
 KW myocardiothrombosis; diagnosis; therapy.

XX Homo sapiens.

XX

PN WO9854199-A1.

XX 03-DEC-1998.

XX 27-MAY-1998; 98WO-US10728.

XX 28-MAY-1997; 97US-0048000.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Ebner R, Moore PA, Ruben SM;

XX WPI; 1999-070207/06.

XX New tissue plasminogen activator-like protease - useful in the  
 PT diagnosis and treatment of circulatory system-related disorders  
 PS Disclosure; Page 57-58; 76pp; English.

XX This is the amino acid sequence of the translation product of the  
 CC human mRNA for tissue plasminogen activator (tPA). It shares  
 CC homology with a novel human polypeptide (see AAW87769), designated  
 CC tissue plasminogen activator-like protease (t-PALP). The homology  
 CC between t-PALP and tPA indicates that t-PALP may be involved in  
 CC the regulation of normal and abnormal clotting in e.g. stroke,  
 CC deep-vein thrombosis, peripheral arterial occlusion, pulmonary  
 CC embolism and myocardiothrombosis.

XX Sequence 372 AA;

Query Match 95.1%; Score 1961; DB 20; Length 372;  
 Best Local Similarity 99.7%; Pred. No. 4.2e-159;  
 Matches 355; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 21 AASEGNSDCYFGNGSAVRGTHSLTESGASCLPWNMSMLIGKVTYAQNPSAQAALGLGKHN 80  
 DB 17 ACSEGNSDCYFGNGSAVRGTHSLTESGASCLPWNMSMLIGKVTYAQNPSAQAALGLGKHN 76  
 QY 81 CRNPDGAKPWCHVKNRRLTWECYDVPSCSTCGLROYSQPOFRIGKGLFADIASHPWQA 140  
 DB 77 CRNPDGAKPWCHVKNRRLTWECYDVPSCSTCGLROYSQPOFRIGKGLFADIASHPWQA 136  
 QY 141 AIFAKHRRSPGERFLCGGILISSCWILSAHCFQERFPPHLLTVILGRTYRVVPEEEOKE 200  
 DB 137 AIFAKHRRSPGERFLCGGILISSCWILSAHCFQERFPPHLLTVILGRTYRVVPEEEOKE 196  
 QY 201 FEVEKIVHKEFDDTDYNDIALQLKSDSRCAQESSVVRTVCLPPADLQLPDWTECEL 260  
 DB 197 FEVEKIVHKEFDDTDYNDIALQLKSDSRCAQESSVVRTVCLPPADLQLPDWTECEL 256  
 QY 261 SGYKHEALSPFYSERLKEAHLVLPSSRCTSOHLNRTVTDNMLCAGDTRSGGPQANLH 320  
 DB 257 SGYKHEALSPFYSERLKEAHLVLPSSRCTSOHLNRTVTDNMLCAGDTRSGGPQANLH 316  
 QY 321 DACQDGGPLVCLNDGRMTLVGIIISWGLGCGQKDVPGVYTKVTNYLDWIRDNRNP 376  
 DB 317 DACQDGGPLVCLNDGRMTLVGIIISWGLGCGQKDVPGVYTKVTNYLDWIRDNRNP 372

RESULT 3

AAW875190

ID AAW875190 standard; Protein; 396 AA.

XX AAW875190;

XX 25-MAR-2003 (updated)

XX 11-DEC-1992 (first entry)

XX mt-PA6.

XX trp; operon; tPA; finger; Kringle domain; EGF; plasminogen; fibrinogen.

XX Escherichia coli.



XX Human; des-epidermal growth factor homologous plasminogen activator;  
 KW tPA; liver membrane; reduced affinity; EGF homologous; thrombosis;  
 KW thrombolytic; increased half-life.  
 XX

OS Homo sapiens (engineered).

XX Key Location/Qualifiers

FT Misc-difference 1..2 /note= "amino acids 2-89 of tPA have been deleted"

XX US376547-A.

PN 27-DEC-1994.

XX 29-JAN-1988; 88US-0150267.

XX 30-JAN-1987; 87US-0008795.

PR 29-JAN-1988; 88US-0150267.

XX (AMHP ) AMERICAN HOME PROD CORP.

XX Hung PP, Kalyan NK, Lee SL;

XX WPI; 1995-043464/06.

XX New modified plasminogen activator cpds. - having regions removed  
 PT to reduce affinity for liver membranes and increase circulation  
 PT half-life.

PS Claim 1; Page ?; 26pp; English.

XX Amino acid residues 2-89 contain the fibronectin and EGF regions of  
 CC human tPA. Deletion of these regions results in a tissue plasminogen  
 CC activator with reduced affinity for liver cell membranes; the  
 CC mutant protein is not cleared from the circulation as rapidly as is  
 CC wild-type tPA. The specification only gives the sequence around the  
 CC deletion and not the full-length sequence of "delta 2-89 tPA"; the  
 CC sequence in AAR6851 has been obtained by amending a previously  
 CC disclosed wild-type human tPA sequence (from DE3930099) according to  
 CC the description given in Example 1.  
 CC (Updated on 25-MAR-2003 to correct PF field.)

XX Sequence 439 AA;

Query Match 95.1%; Score 1961; DB 16; Length 439;  
 Best Local Similarity 99.7%; Pred. No. 5.2e-159;  
 Matches 355; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy	21	AASEGNSDCYFGNGSAVRGTHSLTESGASCLPWNMILIGKVYTAQNPSAALGLGKHNY	80
Db	84	ACSEGNSDCYFGNGSAVRGTHSLTESGASCLPWNMILIGKVYTAQNPSAALGLGKHNY	143
Qy	81	CRNPDGAKPWCHVLKNRRLTWYECVDPSCSTGLROYSQPFRIGKGLFADTASHPWQA	140
Db	144	CRNPDGAKPWCHVLKNRRLTWYECVDPSCSTGLROYSQPFRIGKGLFADTASHPWQA	203
Qy	141	AIFAKHRRSPGERFLCGGILISSCWILSAHCFQERPPPHLTVILGRTYRVVPGEEQK	200
Db	204	AIFAKHRRSPGERFLCGGILISSCWILSAHCFQERPPPHLTVILGRTYRVVPGEEQK	263
Qy	201	FEVEKIIVHKEFDDTDYNDIALQLKSDSRCAQESSVVRTVCLPPADLQPDWTECEL	260
Db	264	FEVEKIIVHKEFDDTDYNDIALQLKSDSRCAQESSVVRTVCLPPADLQPDWTECEL	323
Qy	261	SGYKHEALSPFYSERLKEAHLVLPSSRCTSOHLNRTVTDNMLCAGDTRSGGPOANLH	320
Db	324	SGYKHEALSPFYSERLKEAHLVLPSSRCTSOHLNRTVTDNMLCAGDTRSGGPOANLH	383
Qy	321	DACQDGGPLVCLNDGRMTLVGIIISWGLGCGKQKVPVGYTKVTNYLDWIRDNRMP	376
Db	384	DACQDGGPLVCLNDGRMTLVGIIISWGLGCGKQKVPVGYTKVTNYLDWIRDNRMP	439

# RESULT 6

AAR08150  
 ID AAR08150 standard; protein; 446 AA.

XX AAR08150;

XX 09-JAN-2003 (updated)  
 DT 01-MAR-1991 (first entry)

XX Non-glycosylated tPA deriv. lacking finger- and EGF-domains.  
 DE Tissue plasminogen activator; thrombolysis; finger domain;  
 KW EGF domain.

XX Synthetic.

XX EP400545-A.

XX 05-DEC-1990.

XX 28-MAY-1990; 90EP-0110096.

XX 14-JUL-1989; 89DE-3923339.

PR 31-MAY-1989; 89DE-3917781.

XX (BOEP ) BOEHRINGER MANNHEIM GMBH.

XX Stern A, Kohnert U, Rudolph R, Fischer S, Martin U;

XX WPI; 1990-363094/49.

DR N-PSDB; AAQ06762.

XX New non-glycosylated form of tissue plasminogen activator - with  
 PT thrombolytic activity and long plasma life

XX Claim 1; page 15; 21pp; German.

XX This non-glycosylated tPA deriv. lacks the finger- and EGF-domains.  
 CC It has a lower clearance rate (longer half-life) than natural tPA  
 CC while retaining thrombolytic activity and stimulation by fibrin.  
 CC It is produced by recombinant methods and site-specific muta-  
 CC genesis.

CC (Updated on 09-JAN-2003 to add missing OS field.)

XX Sequence 446 AA;

Query Match 95.1%; Score 1961; DB 11; Length 446;  
 Best Local Similarity 99.7%; Pred. No. 5.3e-159;  
 Matches 355; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy	21	AASEGNSDCYFGNGSAVRGTHSLTESGASCLPWNMILIGKVYTAQNPSAALGLGKHNY	80
Db	91	ACSEGNSDCYFGNGSAVRGTHSLTESGASCLPWNMILIGKVYTAQNPSAALGLGKHNY	150
Qy	81	CRNPDGAKPWCHVLKNRRLTWYECVDPSCSTGLROYSQPFRIGKGLFADTASHPWQA	140
Db	151	CRNPDGAKPWCHVLKNRRLTWYECVDPSCSTGLROYSQPFRIGKGLFADTASHPWQA	210
Qy	141	AIFAKHRRSPGERFLCGGILISSCWILSAHCFQERPPPHLTVILGRTYRVVPGEEQK	200
Db	211	AIFAKHRRSPGERFLCGGILISSCWILSAHCFQERPPPHLTVILGRTYRVVPGEEQK	270
Qy	201	FEVEKIIVHKEFDDTDYNDIALQLKSDSRCAQESSVVRTVCLPPADLQPDWTECEL	260
Db	271	FEVEKIIVHKEFDDTDYNDIALQLKSDSRCAQESSVVRTVCLPPADLQPDWTECEL	330
Qy	261	SGYKHEALSPFYSERLKEAHLVLPSSRCTSOHLNRTVTDNMLCAGDTRSGGPOANLH	320
Db	331	SGYKHEALSPFYSERLKEAHLVLPSSRCTSOHLNRTVTDNMLCAGDTRSGGPOANLH	390
Qy	321	DACQDGGPLVCLNDGRMTLVGIIISWGLGCGKQKVPVGYTKVTNYLDWIRDNRMP	376

**THE UNIVERSITY OF CHICAGO**

357 SCYSCHEIAT SCFVSCFV KRAIHPV YVDSFVCTSCVH I MHTMTHMHT CACHTMPCSCCPCPCVW W 41

357 SCYSCHEIAT SCFVSCFV KRAIHPV YVDSFVCTSCVH I MHTMTHMHT CACHTMPCSCCPCPCVW W 41

```
Db 284 AIFAKHRRSPGERFLCGGILISSCWILSAAHCFQERPPPHLTVILGRTYRVVPEEEQK 343
Qy 201 FEVEKYIVHKEFDDTYDNDIALQLKSDSRCAQESSVVRTVCLPPADLQLPDWTCECL 260
Db 344 FEVEKYIVHKEFDDTYDNDIALQLKSDSRCAQESSVVRTVCLPPADLQLPDWTCECL 403
Qy 261 SGYGKEALSPFYSERLKEAHVRLYPSSRCTSQHLLNRTVTDNMLCAGDTRSGGQANLH 320
Db 404 SGYGKEALSPFYSERLKEAHVRLYPSSRCTSQHLLNRTVTDNMLCAGDTRSGGQANLH 463
Qy 321 DACQDSGGPLVCLNDGRMTLVGIIISWGLGCGKQDVPGVYTKVTNYLDWIRDNRMP 376
Db 464 DACQDSGGPLVCLNDGRMTLVGIIISWGLGCGKQDVPGVYTKVTNYLDWIRDNRMP 519

RESULT 9
AAR05489
ID AAR05489 standard; protein; 527 AA.
AC AAR05489;
XX
DT 25-MAR-2003 (updated)
DT 18-OCT-1990 (first entry)
DE tPA024 precursor protein.
KW Tissue plasminogen activator; TPA;fibrin;
KW myocardial infarction; ds.
OS Homo sapiens.
XX
FH Key 86 Location/Qualifiers
FT Modified-site /note="Asn linked to oligosaccharide "
XX
PN EP373896-A.
XX
PD 20-JUN-1990.
XX
PF 12-DEC-1989; 89EP-0312993.
XX
PR 12-DEC-1988; 88JP-0314172.
XX
PA (YAMA ) YAMANOUCHI PHARM CO LTD.
PA (YAMA ) NIPPON STEEL CORP.
XX
PI Shimizu Y, Yano E, Yano S, Kato M, Kinoshita A, Kawasaki T;
PI Ishida J, Gushima H;
XX
WPI; 1990-187265/25.
DR N-PSDB; AAQ04903.
XX
PT New tissue plasminogen activator analogues -
PT contg. glycosylated asparagine unit for increased half-life.
XX
PS Disclosure; Page ?; 34pp; English.
XX
CC The tPA analogue carries an oligosaccharide side chain linked to an
CC Asn residue replacing either Tyr-67 or Ile-86. The nature of the
CC side chain depends on the host cell used.
CC The analogue has a longer half life and may be administered at
CC lower doses.
CC (Updated on 25-MAR-2003 to correct PA field.)
XX
SQ Sequence 527 AA;

Query Match 95.1%; Score 1961; DB 11; Length 527;
Best Local Similarity 99.7%; Pred. No. 6.4e-159;
Matches 355; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 21 AASEGNSDCYFGNGSAYRGTHTSLTESGASCLPWNMILIGKVYTAQNPSAALGLGKHNY 80
Db 172 ACSEGNSDCYFGNGSAYRGTHTSLTESGASCLPWNMILIGKVYTAQNPSAALGLGKHNY 231
Qy 81 CRNPDGDAKPCWCHLVKNRRLTWECYCDVPSCSTGLRQYSDPOFRKIGGLPADIASHPQOA 140
Db 232 CRNPDGDAKPCWCHLVKNRRLTWECYCDVPSCSTGLRQYSDPOFRKIGGLPADIASHPQOA 291
Qy 141 AIFAKHRRSPGERFLCGGILISSCWILSAAHCFQERPPPHLTVILGRTYRVVPEEEQK 200
Db 292 AIFAKHRRSPGERFLCGGILISSCWILSAAHCFQERPPPHLTVILGRTYRVVPEEEQK 351
Qy 201 FEVEKYIVHKEFDDTYDNDIALQLKSDSRCAQESSVVRTVCLPPADLQLPDWTCECL 260
Db 352 FEVEKYIVHKEFDDTYDNDIALQLKSDSRCAQESSVVRTVCLPPADLQLPDWTCECL 411
Qy 261 SGYGKEALSPFYSERLKEAHVRLYPSSRCTSQHLLNRTVTDNMLCAGDTRSGGQANLH 320
Db 412 SGYGKEALSPFYSERLKEAHVRLYPSSRCTSQHLLNRTVTDNMLCAGDTRSGGQANLH 471
Qy 321 DACQDSGGPLVCLNDGRMTLVGIIISWGLGCGKQDVPGVYTKVTNYLDWIRDNRMP 376
Db 472 DACQDSGGPLVCLNDGRMTLVGIIISWGLGCGKQDVPGVYTKVTNYLDWIRDNRMP 527

RESULT 10
AAR05488
ID AAR05488 standard; protein; 527 AA.
XX
AC AAR05488;
XX
DT 25-MAR-2003 (updated)
DT 18-OCT-1990 (first entry)
DE tPA024 precursor protein.
KW Tissue plasminogen activator; TPA;fibrin;
KW myocardial infarction; ds.
OS Homo sapiens.
XX
FH Key 67 Location/Qualifiers
FT Modified-site /note="Asn linked to oligosaccharide"
XX
PN EP373896-A.
XX
PD 20-JUN-1990.
XX
PF 12-DEC-1989; 89EP-0312993.
XX
PR 12-DEC-1988; 88JP-0314172.
XX
PA (YAMA ) YAMANOUCHI PHARM CO LTD.
PA (YAMA ) NIPPON STEEL CORP.
XX
PI Shimizu Y, Yano E, Yano S, Kato M, Kinoshita A, Kawasaki T;
PI Ishida J, Gushima H;
XX
WPI; 1990-187265/25.
DR N-PSDB; AAQ04904.
XX
PT New tissue plasminogen activator analogues -
PT contg. glycosylated asparagine unit for increased half-life.
XX
PS Disclosure; Page ?; 34pp; English.
XX
CC The tPA analogue carries an oligosaccharide side chain linked to an
CC Asn residue replacing either Tyr-67 or Ile-86. The nature of the
CC side chain depends on the host cell used.
CC The analogue has a longer half life and may be administered at
CC lower doses.
CC (Updated on 25-MAR-2003 to correct PA field.)
XX
SQ Sequence 527 AA;
```



Query Match 95.1%; Score 1961; DB 11; Length 527;  
 Best Local Similarity 99.7%; Pred. No. 6.4e-159;  
 Matches 355; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 21 AASEGSDCYFGNGSAYRGTHSLTESGASCLPWNMSMLIGKVTYTAQNPSPAALGLGKHNY 80  
 DB 172 ACSEGSDCYFGNGSAYRGTHSLTESGASCLPWNMSMLIGKVTYTAQNPSPAALGLGKHNY 231

QY 81 CRNPDGDAKPCWCHVKNRRLTWECYDVPSCSTCGLROYSQPFRIKGGLFADIASHPWOA 140  
 DB 232 CRNPDGDAKPCWCHVKNRRLTWECYDVPSCSTCGLROYSQPFRIKGGLFADIASHPWOA 291

QY 141 AIFAKHRRSPGERFLCGGILISSCWILSAAHCFQERPPPHLTVILGRTYRVVPEBEQK 200  
 DB 292 AIFAKHRRSPGERFLCGGILISSCWILSAAHCFQERPPPHLTVILGRTYRVVPEBEQK 351

QY 201 FEVEKYIVHKEFDDDTYDNDIALQLKSDSRCAQESSVVRTVCLPPADLQLPDWTCECEL 260  
 DB 352 FEVEKYIVHKEFDDDTYDNDIALQLKSDSRCAQESSVVRTVCLPPADLQLPDWTCECEL 411

QY 261 SGYKGHEALSPFYSERLKEAHRVLYPSSRCTSOHLLNRTVTDNMLCAGDTRSGGPOANLH 320  
 DB 412 SGYKGHEALSPFYSERLKEAHRVLYPSSRCTSOHLLNRTVTDNMLCAGDTRSGGPOANLH 471

QY 321 DACQSDSGPLVCLNDGRMTLVGIISWGLGCGKQDVPGVYTKVTNYLDWIRDNRMP 376  
 DB 472 DACQSDSGPLVCLNDGRMTLVGIISWGLGCGKQDVPGVYTKVTNYLDWIRDNRMP 527

## RESULT 11

AAR13910  
 ID AAR13910 standard; protein; 527 AA.  
 AC AAR13910;  
 DT 25-MAR-2003 (updated)  
 DT 25-NOV-1991 (first entry)  
 DE T-PA deriv. (I).  
 KW Tissue plasminogen activator; thrombolysis; myocardial infarction.  
 OS Homo sapiens.  
 PN EP445464-A.  
 PD 11-SEP-1991.  
 PX 15-NOV-1990; 90EP-0312445.  
 PX 15-NOV-1990; 90EP-0312445.  
 PR 01-MAR-1990; 90JP-0050428.  
 PR 02-AUG-1990; 90JP-0206458.  
 PX (KANF ) KANEGAFUCHI KAGAKU KOGYO KK.  
 PX Yahara H, Nagaoka T, Yajima K, Ikenaka Y, Matsumoto K;  
 PI Kakutani T;  
 DR WPI; 1991-268391/37.  
 XX New tissue plasminogen derivs. - having amino acid replacements  
 PT to increase persistence in the blood while retaining thrombolytic  
 PT activity  
 XX Claim 4; Page 29; 39pp; English.  
 XX The t-PA has the following amino acids of the mature protein  
 CC replaced: N(37) for S, S(38) for V, G(39) for V, R(40) for E,  
 CC A(41) for S and Q(42) for S. The deriv. has improved persistence in  
 CC the blood compared with native t-PA, while retaining the same level of  
 CC thrombolytic activity. It can be used for the treatment of

CC thrombotic disorders such as myocardial infarction.  
 CC The wild-type t-PA sequence was retrieved by the indexer  
 CC (DE3930099) and "mutated" according to the claims.  
 CC See also AAR13910-14.  
 CC (Updated on 25-MAR-2003 to correct PA field.)  
 XX  
 SQ Sequence 527 AA;

Query Match 95.1%; Score 1961; DB 12; Length 527;  
 Best Local Similarity 99.7%; Pred. No. 6.4e-159;  
 Matches 355; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 21 AASEGSDCYFGNGSAYRGTHSLTESGASCLPWNMSMLIGKVTYTAQNPSPAALGLGKHNY 80  
 DB 172 ACSEGSDCYFGNGSAYRGTHSLTESGASCLPWNMSMLIGKVTYTAQNPSPAALGLGKHNY 231

QY 81 CRNPDGDAKPCWCHVKNRRLTWECYDVPSCSTCGLROYSQPFRIKGGLFADIASHPWOA 140  
 DB 232 CRNPDGDAKPCWCHVKNRRLTWECYDVPSCSTCGLROYSQPFRIKGGLFADIASHPWOA 291

QY 141 AIFAKHRRSPGERFLCGGILISSCWILSAAHCFQERPPPHLTVILGRTYRVVPEBEQK 200  
 DB 292 AIFAKHRRSPGERFLCGGILISSCWILSAAHCFQERPPPHLTVILGRTYRVVPEBEQK 351

QY 201 FEVEKYIVHKEFDDDTYDNDIALQLKSDSRCAQESSVVRTVCLPPADLQLPDWTCECEL 260  
 DB 352 FEVEKYIVHKEFDDDTYDNDIALQLKSDSRCAQESSVVRTVCLPPADLQLPDWTCECEL 411

QY 261 SGYKGHEALSPFYSERLKEAHRVLYPSSRCTSOHLLNRTVTDNMLCAGDTRSGGPOANLH 320  
 DB 412 SGYKGHEALSPFYSERLKEAHRVLYPSSRCTSOHLLNRTVTDNMLCAGDTRSGGPOANLH 471

QY 321 DACQSDSGPLVCLNDGRMTLVGIISWGLGCGKQDVPGVYTKVTNYLDWIRDNRMP 376  
 DB 472 DACQSDSGPLVCLNDGRMTLVGIISWGLGCGKQDVPGVYTKVTNYLDWIRDNRMP 527

## RESULT 12

AAR13911  
 ID AAR13911 standard; protein; 527 AA.  
 AC AAR13911;  
 DT 25-MAR-2003 (updated)  
 DT 25-NOV-1991 (first entry)  
 DE T-PA deriv. (II).  
 KW Tissue plasminogen activator; thrombolysis; myocardial infarction.  
 OS Homo sapiens.  
 PN EP445464-A.  
 PD 11-SEP-1991.  
 PX 15-NOV-1990; 90EP-0312445.  
 PX 15-NOV-1990; 90EP-0312445.  
 PR 01-MAR-1990; 90JP-0050428.  
 PR 02-AUG-1990; 90JP-0206458.  
 PX (KANF ) KANEGAFUCHI KAGAKU KOGYO KK.  
 PX Yahara H, Nagaoka T, Yajima K, Ikenaka Y, Matsumoto K;  
 PI Kakutani T;  
 DR WPI; 1991-268391/37.  
 XX New tissue plasminogen derivs. - having amino acid replacements  
 PT to increase persistence in the blood while retaining thrombolytic  
 PT activity  
 XX

PS Claim 5; Page 29; 39pp; English.

XX The t-PA has the following amino acids of the mature protein  
XX replaced: N(37) for S, S(38) for V, G(39) for V, R(40) for E,  
CC A(41) for F and Q(42) for S. The deriv. has improved persistence in  
CC the blood compared with native t-PA, while retaining the same level of  
CC thrombolytic activity. It can be used for the treatment of  
CC thrombotic disorders such as myocardial infarction.  
CC The wild-type t-PA sequence was retrieved by the indexer  
CC (DE3930099) and "mutated" according to the claims.  
CC See also AAR13910-14.  
CC (Updated on 25-MAR-2003 to correct PA field.)

SQ Sequence 527 AA;

Query Match 95.1%; Score 1961; DB 12; Length 527;  
Best Local Similarity 99.7%; Pred. No. 6.4e-159;  
Matches 355; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 21 AASEGNSDCYFGNGSAVRGTHSLTESGASCLPWNMILIGKVYTAQNPSAQLGLGKHNY 80  
DB 172 ACSEGNSDCYFGNGSAVRGTHSLTESGASCLPWNMILIGKVYTAQNPSAQLGLGKHNY 231  
QY 81 CRNPDGDAKPCWCHLVKNRRLTWECYDVPSCSTGLRQYSPQFRIKGLFADIAHPWQA 140  
DB 232 CRNPDGDAKPCWCHLVKNRRLTWECYDVPSCSTGLRQYSPQFRIKGLFADIAHPWQA 291  
QY 141 AIFAKHRRSPGERFLCGGILISSCWILSAHCFQERPPPHLTVILGRTYRVVPEBEQK 200  
DB 292 AIFAKHRRSPGERFLCGGILISSCWILSAHCFQERPPPHLTVILGRTYRVVPEBEQK 351  
QY 201 FEVEKYIVHKEFDDDTYDNDIALQLKSDSRCAQESSVVRTVCLPPADLPDWTCECL 260  
DB 352 FEVEKYIVHKEFDDDTYDNDIALQLKSDSRCAQESSVVRTVCLPPADLPDWTCECL 411  
QY 261 SGYGKHEALSPFYSERLKEAHRVLPSSRCTSQHLLNRTVTDNMLCAGDTRSGGPQANLH 320  
DB 412 SGYGKHEALSPFYSERLKEAHRVLPSSRCTSQHLLNRTVTDNMLCAGDTRSGGPQANLH 471  
QY 321 DACQSGDGPLVCLNDGRMTLVGIIISWGLCGCKQKDVPGVYTKVTNYLDWIRDNRPP 376  
DB 472 DACQSGDGPLVCLNDGRMTLVGIIISWGLCGCKQKDVPGVYTKVTNYLDWIRDNRPP 527

RESULT 13

AAR13912  
ID AAR13912 standard; Protein; 527 AA.

XX AC AAR13912;

DT 25-MAR-2003 (updated)

DT 25-NOV-1991 (first entry)

DE T-PA deriv. (III).

XX Tissue plasminogen activator; thrombolysis; myocardial infarction.

XX Homo sapiens.

XX EP445464-A.

XX 11-SEP-1991.

XX 15-NOV-1990; 90EP-0312445.

XX 15-NOV-1990; 90EP-0312445.

XX 01-MAR-1990; 90JP-0050428.

XX 02-AUG-1990; 90JP-0206458.

XX (KANF) KANEKAFUCHI KAGAKU KOGYO KK.

XX Yahara H, Nagaoka T, Yajima K, Ikenaka Y, Matsumoto K;

PI Kakutani T;

XX WPI; 1991-268391/37.

XX New tissue plasminogen derivs. - having amino acid replacements  
PT to increase persistence in the blood while retaining thrombolytic  
PT activity

PS Claim 6; Page 29; 39pp; English.

XX The t-PA has the following amino acids of the mature protein  
CC replaced: N(37) for S, S(38) for V, G(39) for V, R(40) for I,  
CC A(41) for V and Q(42) for S. The deriv. has improved persistence in  
CC the blood compared with native t-PA, while retaining the same level of  
CC thrombolytic activity. It can be used for the treatment of  
CC thrombotic disorders such as myocardial infarction.  
CC The wild-type t-PA sequence was retrieved by the indexer  
CC (DE3930099) and "mutated" according to the claims.  
CC See also AAR13910-14.  
CC (Updated on 25-MAR-2003 to correct PA field.)

SQ Sequence 527 AA;

Query Match 95.1%; Score 1961; DB 12; Length 527;  
Best Local Similarity 99.7%; Pred. No. 6.4e-159;  
Matches 355; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 21 AASEGNSDCYFGNGSAVRGTHSLTESGASCLPWNMILIGKVYTAQNPSAQLGLGKHNY 80  
DB 172 ACSEGNSDCYFGNGSAVRGTHSLTESGASCLPWNMILIGKVYTAQNPSAQLGLGKHNY 231  
QY 81 CRNPDGDAKPCWCHLVKNRRLTWECYDVPSCSTGLRQYSPQFRIKGLFADIAHPWQA 140  
DB 232 CRNPDGDAKPCWCHLVKNRRLTWECYDVPSCSTGLRQYSPQFRIKGLFADIAHPWQA 291  
QY 141 AIFAKHRRSPGERFLCGGILISSCWILSAHCFQERPPPHLTVILGRTYRVVPEBEQK 200  
DB 292 AIFAKHRRSPGERFLCGGILISSCWILSAHCFQERPPPHLTVILGRTYRVVPEBEQK 351  
QY 201 FEVEKYIVHKEFDDDTYDNDIALQLKSDSRCAQESSVVRTVCLPPADLPDWTCECL 260  
DB 352 FEVEKYIVHKEFDDDTYDNDIALQLKSDSRCAQESSVVRTVCLPPADLPDWTCECL 411  
QY 261 SGYGKHEALSPFYSERLKEAHRVLPSSRCTSQHLLNRTVTDNMLCAGDTRSGGPQANLH 320  
DB 412 SGYGKHEALSPFYSERLKEAHRVLPSSRCTSQHLLNRTVTDNMLCAGDTRSGGPQANLH 471  
QY 321 DACQSGDGPLVCLNDGRMTLVGIIISWGLCGCKQKDVPGVYTKVTNYLDWIRDNRPP 376  
DB 472 DACQSGDGPLVCLNDGRMTLVGIIISWGLCGCKQKDVPGVYTKVTNYLDWIRDNRPP 527

RESULT 14

AAR13913  
ID AAR13913 standard; Protein; 527 AA.

XX AC AAR13913;

DT 25-MAR-2003 (updated)

DT 25-NOV-1991 (first entry)

XX T-PA deriv. (IV).

XX Tissue plasminogen activator; thrombolysis; myocardial infarction.

XX Homo sapiens.

XX EP445464-A.

XX 11-SEP-1991.

XX 15-NOV-1990; 90EP-0312445.

XX 15-NOV-1990; 90EP-0312445.

```

PR 01-MAR-1990; 90JP-0050428.
XX 02-AUG-1990; 90JP-0206458.
XX (KANF ) KANEGAFUCHI KAGAKU KOGYO KK.
XX Yahara H, Nagaoka T, Yajima K, Ikenaka Y, Matsumoto K;
XX Kakutani T;
XX WPI; 1991-268391/37.
XX New tissue plasminogen derivs. - having amino acid replacements
XX to increase persistence in the blood while retaining thrombolytic
XX activity
XX Claim 7; Page 29; 39pp; English.
XX The t-PA has the following amino acids of the mature protein
XX replaced: G(161) for R, K(162) for R and S(165) for W.
XX The deriv. has improved persistence in the blood compared with native
XX t-PA, while retaining the same level of thrombolytic activity.
XX It can be used for the treatment of thrombotic disorders such as
XX myocardial infarction.
XX The wild-type t-PA sequence was retrieved by the indexer
XX (DE3930099) and "mutated" according to the claims.
XX See also AAR13910-14.
XX (Updated on 25-MAR-2003 to correct PA field.)
XX SQ Sequence 527 AA;

Query Match .95.1%; Score 1961; DB 12; Length 527;
Best Local Similarity 99.7%; Pred. No. 6.4e-159;
Matches 355; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 21 AASEGNSDCYFGNGSAYRGTHSLTESGASCLPWNMILIGKVYTAQNPSAALGLGHNY 80
DB 172 ACSEGNSDCYFGNGSAYRGTHSLTESGASCLPWNMILIGKVYTAQNPSAALGLGHNY 231

QY 81 CRNPDGDAKPCWCHLVKNRRLTWECYDVPCSTCGLRQYSPQPRKIGGLFADIASHPWQA 140
DB 232 CRNPDGDAKPCWCHLVKNRRLTWECYDVPCSTCGLRQYSPQPRKIGGLFADIASHPWQA 291

QY 141 AIFAKHRRSPGERFLCGGILISSCWILSAACHQFERPPPHLTVILGRTYRVVPGEEOK 200
DB 292 AIFAKHRRSPGERFLCGGILISSCWILSAACHQFERPPPHLTVILGRTYRVVPGEEOK 351

QY 201 FEVEKYIVHKEFDDDTYDNDIALQLKSDSRCAQESSVVRTVCLPPADLQPDWTECEL 260
DB 352 FEVEKYIVHKEFDDDTYDNDIALQLKSDSRCAQESSVVRTVCLPPADLQPDWTECEL 411

QY 261 SGYKHEALSPFYSERLKEAHRVLYPSSRCTSOHLNRTVTDNMLCAGDTRSGGQPANLH 320
DB 412 SGYKHEALSPFYSERLKEAHRVLYPSSRCTSOHLNRTVTDNMLCAGDTRSGGQPANLH 471

QY 321 DACQDSGGPLVCLNDGRMTLVGIISWGLGCGQKDFGVYTKVTNYLDWIRDNRNP 376
DB 472 DACQDSGGPLVCLNDGRMTLVGIISWGLGCGQKDFGVYTKVTNYLDWIRDNRNP 527

RESULT 15
AAR13914
ID AAR13914 standard; Protein; 527 AA.
XX AC AAR13914;
XX DT 25-MAR-2003 (updated)
XX DT 25-NOV-1991 (first entry)
XX DE T-PA deriv. (V).
XX KW Tissue plasminogen activator; thrombolysis; myocardial infarction.
XX OS Homo sapiens.
XX

```

---

```

FH Key Location/Qualifiers
FT Misc-difference 37
FT /label= ASN, SER
FT Misc-difference 38
FT /label= SER, VAL
FT Misc-difference 39
FT /label= GLY, VAL
FT Misc-difference 40
FT /label= ARG, GLU
FT Misc-difference 41
FT /label= ALA, SER
FT Misc-difference 42
FT /label= GLN, SER
XX EP445464-A.
XX 11-SEP-1991.
XX 15-NOV-1990; 90EP-0312445.
XX 15-NOV-1990; 90EP-0312445.
XX 01-MAR-1990; 90JP-0050428.
XX 02-AUG-1990; 90JP-0206458.
XX (KANF ) KANEGAFUCHI KAGAKU KOGYO KK.
XX Yahara H, Nagaoka T, Yajima K, Ikenaka Y, Matsumoto K;
XX Kakutani T;
XX WPI; 1991-268391/37.
XX New tissue plasminogen derivs. - having amino acid replacements
XX to increase persistence in the blood while retaining thrombolytic
XX activity
XX Claim 10; Page 29; 39pp; English.
XX The t-PA has the following amino acids of the mature protein
XX replaced: N(115) for P and one of N(37) for S, S(38) for V, G(39)
XX for V, R(40) for E, A(41) for S and Q(42) for S.
XX The deriv. has improved persistence in the blood compared with native
XX t-PA, while retaining the same level of thrombolytic activity.
XX It can be used for the treatment of thrombotic disorders such as
XX myocardial infarction.
XX The wild-type t-PA sequence was retrieved by the indexer
XX (DE3930099) and "mutated" according to the claims.
XX See also AAR13910-14.
XX (Updated on 25-MAR-2003 to correct PA field.)
XX SQ Sequence 527 AA;

Query Match .95.1%; Score 1961; DB 12; Length 527;
Best Local Similarity 99.7%; Pred. No. 6.4e-159;
Matches 355; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 21 AASEGNSDCYFGNGSAYRGTHSLTESGASCLPWNMILIGKVYTAQNPSAALGLGHNY 80
DB 172 ACSEGNSDCYFGNGSAYRGTHSLTESGASCLPWNMILIGKVYTAQNPSAALGLGHNY 231

QY 81 CRNPDGDAKPCWCHLVKNRRLTWECYDVPCSTCGLRQYSPQPRKIGGLFADIASHPWQA 140
DB 232 CRNPDGDAKPCWCHLVKNRRLTWECYDVPCSTCGLRQYSPQPRKIGGLFADIASHPWQA 291

QY 141 AIFAKHRRSPGERFLCGGILISSCWILSAACHQFERPPPHLTVILGRTYRVVPGEEOK 200
DB 292 AIFAKHRRSPGERFLCGGILISSCWILSAACHQFERPPPHLTVILGRTYRVVPGEEOK 351

QY 201 FEVEKYIVHKEFDDDTYDNDIALQLKSDSRCAQESSVVRTVCLPPADLQPDWTECEL 260
DB 352 FEVEKYIVHKEFDDDTYDNDIALQLKSDSRCAQESSVVRTVCLPPADLQPDWTECEL 411

QY 261 SGYKHEALSPFYSERLKEAHRVLYPSSRCTSOHLNRTVTDNMLCAGDTRSGGQPANLH 320

```

Db 412 SGGKHEALSPFYSERLKEAHVRLYPSSRCTSQHLLNRTVTDNMLCAGDTRSGGPQANLH 471  
Qy 321 DACOGDSGGPLVCLNDGRMTLVGIISWGLGCGQXDVPGVYTKVTNYLDWIRDNMRP 376  
Db 472 DACOGDSGGPLVCLNDGRMTLVGIISWGLGCGQXDVPGVYTKVTNYLDWIRDNMRP 527

Search completed: August 8, 2003, 17:13:51  
Job time : 45 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 8, 2003, 17:13:07 ; Search time 20 Seconds  
(without alignments)  
797.560 Million cell updates/sec

Title: US-09-987-455-8

Perfect score: 2061

Sequence: 1 MKKTAIAIALAGFATVAQ.....GVYTKVTNYLDWIRDNRRPG 377

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.\*

- 1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep.\*
- 2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep.\*
- 3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep.\*
- 4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep.\*
- 5: /cgn2\_6/ptodata/1/iaa/PTCUS\_COMB.pep.\*
- 6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1961	95.1	472	2	US-08-811-949-63
2	1961	95.1	527	1	US-07-609-510B-16
3	1961	95.1	527	5	PCT-US91-01025A-2
4	1961	95.1	527	6	5185259-8
5	1961	95.1	562	2	US-08-811-949-43
6	1961	95.1	562	2	US-08-560-098A-50
7	1961	95.1	562	2	US-08-883-795A-38
8	1961	95.1	562	6	5185259-3
9	1961	95.1	562	6	5200340-2
10	1961	95.1	562	6	5344773-2
11	1958	95.0	562	6	5244676-5
12	1957	95.0	355	2	US-08-811-949-45
13	1957	95.0	437	2	US-08-811-949-49
14	1955.5	94.9	389	2	US-08-811-949-65
15	1954	94.8	437	2	US-08-811-949-55
16	1953	94.8	437	2	US-08-811-949-51
17	1953	94.8	527	2	US-08-811-949-39
18	1951.5	94.7	378	4	US-09-553-498-10
19	1951.5	94.7	378	4	US-09-618-869-10
20	1950	94.6	355	1	US-08-137-116-1
21	1950	94.6	355	1	US-08-217-618-1
22	1950	94.6	355	1	US-08-427-640-2
23	1950	94.6	355	1	US-08-217-617A-1
24	1950	94.6	355	1	US-08-217-616-1
25	1950	94.6	355	2	US-08-811-949-53
26	1950	94.6	355	3	US-08-794-528-1
27	1950	94.6	355	6	5223256-1

28 1949 94.6 355 2 US-08-811-949-47 Sequence 47, Appl  
29 1946 94.4 437 2 US-08-811-949-57 Sequence 57, Appl  
30 1943.5 94.3 389 2 US-08-811-949-67 Sequence 67, Appl  
31 1935 93.9 355 1 US-08-427-640-6 Sequence 6, Appl  
32 1933.5 93.8 356 1 US-08-427-640-4 Sequence 4, Appl  
33 1931 93.7 355 2 US-08-811-949-59 Sequence 59, Appl  
34 1929 93.6 527 6 5520913-1 Patent No. 5520913  
35 1919.5 93.1 354 2 US-08-811-949-61 Sequence 61, Appl  
36 1882.5 91.3 347 2 US-08-811-949-1 Sequence 1, Appl  
37 1787 86.7 356 1 US-08-427-640-8 Sequence 8, Appl  
38 1704 82.7 326 4 US-09-411-977-3 Sequence 3, Appl  
39 1458 70.7 383 2 US-08-558-269-6 Sequence 6, Appl  
40 1458 70.7 383 3 US-09-410-882-6 Sequence 6, Appl  
41 1423 69.0 477 2 US-08-560-098A-51 Sequence 51, Appl  
42 1380 67.0 253 2 US-09-027-337-8 Sequence 8, Appl  
43 1380 67.0 253 4 US-09-644-600-8 Sequence 8, Appl  
44 1375 66.7 252 3 US-08-944-483-72 Sequence 72, Appl  
45 880.5 42.7 546 6 5200340-6 Patent No. 5200340

#### ALIGNMENTS

RESULT 1  
US-08-811-949-63  
; Sequence 63, Application US/08811949  
; Patent No. 5840533  
; GENERAL INFORMATION:  
; APPLICANT: NIWA, MINO  
; APPLICANT: SAITO, YOSHIMASA  
; APPLICANT: SASAKI, HITOSHI  
; APPLICANT: HAYASHI, MASAKO  
; APPLICANT: NOTANI, JOUJI  
; APPLICANT: KOBAYASHI, MASAKAZU  
; TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR  
; NUMBER OF SEQUENCES: 67  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,  
; ADDRESSEE: P.C.  
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400  
; CITY: ARLINGTON  
; STATE: VA  
; COUNTRY: USA  
; ZIP: 22202  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/811,949  
; FILING DATE: 05-MAR-1997  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: OBLON, NORMAN F.  
; REGISTRATION NUMBER: 24,618  
; REFERENCE/DOCKET NUMBER: 18-966-0  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 703-413-3000  
; TELEFAX: 703-413-2220  
; INFORMATION FOR SEQ ID NO: 63:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 472 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-811-949-63

Query Match 95.1%; Score 1961; DB 2; Length 472;  
Best Local Similarity 99.7%; Pred. No. 8e-186;  
Matches 355; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 21 AASEGNSDCYFGNGSAVRGTHSLTSGASCLPWNMSILGKVVYTAQNPASQALGLGKHNY 80

Db 117 ACSEGNDCYFCGNSAVRGTHSLTESGASCLPWNMILIGKVTYTAQNPSAALGLGKHNY 176  
Qy 81 CRNPDGDAKPNCHVLKNRRLTWECYCDVPCSTCGLRQYSDQFRIKGLFADIAHPWQA 140  
Db 177 CRNPDGDAKPNCHVLKNRRLTWECYCDVPCSTCGLRQYSDQFRIKGLFADIAHPWQA 236  
Qy 141 AIFAKHRRSPGERFLCGGILISSCWILSAACHFCQERPPPHLTVILGRTYRVVPGEEQK 200  
Db 237 AIFAKHRRSPGERFLCGGILISSCWILSAACHFCQERPPPHLTVILGRTYRVVPGEEQK 296  
Qy 201 FEVEKYIVHKEFDDDDTYDNDIALQLKSDSSRCAQESSVVRTVCLPPADLQLPDWTCECL 260  
Db 297 FEVEKYIVHKEFDDDDTYDNDIALQLKSDSSRCAQESSVVRTVCLPPADLQLPDWTCECL 356  
Qy 261 SYGKHEALSPFYSERLKEAHRVLYPSSRCTSOHLNRTVTDNMLCAGDTRSGGQANLH 320  
Db 357 SYGKHEALSPFYSERLKEAHRVLYPSSRCTSOHLNRTVTDNMLCAGDTRSGGQANLH 416  
Qy 321 DACQSDSGPLVCLNDGRMTLVGIIISWGLGCGKQDVPGVYTKVTNYLDWIRDNNRP 376  
Db 417 DACQSDSGPLVCLNDGRMTLVGIIISWGLGCGKQDVPGVYTKVTNYLDWIRDNNRP 472

## RESULT 2

US-07-609-510B-16  
; Sequence 16, Application US/07609510B  
; Patent No. 5326700  
; GENERAL INFORMATION:  
; APPLICANT: Berg et al.  
; TITLE OF INVENTION: Method for Altering Post-Translational Processing of Tissue P  
; NUMBER OF SEQUENCES: 16  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Eli Lilly and Company  
; STREET: Lilly Corporate Center  
; CITY: Indianapolis  
; STATE: IN.  
; COUNTRY: U.S.A.  
; ZIP: 46285  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.0 Mb storage  
; COMPUTER: Macintosh  
; OPERATING SYSTEM: Macintosh  
; SOFTWARE: Microsoft Word  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07609,510B  
; FILING DATE: 19901106  
; CLASSIFICATION: 435  
; INFORMATION FOR SEQ ID NO: 16:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 527 amino acids  
; TYPE: AMINO ACID  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-07-609-510B-16

Query Match 95.1%; Score 1961; DB 1; Length 527;  
Best Local Similarity 99.7%; Pred. No. 9.3e-186;  
Matches 355; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 21 AASEGNSDCYFCGNSAVRGTHSLTESGASCLPWNMILIGKVTYTAQNPSAALGLGKHNY 80  
Db 172 ACSEGNDCYFCGNSAVRGTHSLTESGASCLPWNMILIGKVTYTAQNPSAALGLGKHNY 231  
Qy 81 CRNPDGDAKPNCHVLKNRRLTWECYCDVPCSTCGLRQYSDQFRIKGLFADIAHPWQA 140  
Db 232 CRNPDGDAKPNCHVLKNRRLTWECYCDVPCSTCGLRQYSDQFRIKGLFADIAHPWQA 291  
Qy 141 AIFAKHRRSPGERFLCGGILISSCWILSAACHFCQERPPPHLTVILGRTYRVVPGEEQK 200  
Db 292 AIFAKHRRSPGERFLCGGILISSCWILSAACHFCQERPPPHLTVILGRTYRVVPGEEQK 351

Qy 201 FEVEKYIVHKEFDDDDTYDNDIALQLKSDSSRCAQESSVVRTVCLPPADLQLPDWTCECL 260  
Db 352 FEVEKYIVHKEFDDDDTYDNDIALQLKSDSSRCAQESSVVRTVCLPPADLQLPDWTCECL 411  
Qy 261 SYGKHEALSPFYSERLKEAHRVLYPSSRCTSOHLNRTVTDNMLCAGDTRSGGQANLH 320  
Db 412 SYGKHEALSPFYSERLKEAHRVLYPSSRCTSOHLNRTVTDNMLCAGDTRSGGQANLH 471  
Qy 321 DACQSDSGPLVCLNDGRMTLVGIIISWGLGCGKQDVPGVYTKVTNYLDWIRDNNRP 376  
Db 472 DACQSDSGPLVCLNDGRMTLVGIIISWGLGCGKQDVPGVYTKVTNYLDWIRDNNRP 527

## RESULT 3

PCT-US91-01025A-2  
; Sequence 2, Application PC/TUS9101025A  
; GENERAL INFORMATION:  
; APPLICANT: Genentech, Inc.  
; TITLE OF INVENTION: Tissue Plasminogen Activator Having Fibrin  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Genentech, Inc.  
; STREET: 460 Point San Bruno Blvd  
; CITY: South San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94080  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: patin (Genentech)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US91/01025A  
; FILING DATE: 19910214  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/486,657  
; FILING DATE: 1 March 1990  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Hasek, Janet E.  
; REGISTRATION NUMBER: 28,616  
; REFERENCE/DOCKET NUMBER: 454P2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415/266-1896  
; TELEFAX: 415/952-9881  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 527 amino acids  
; TYPE: AMINO ACID  
; TOPOLOGY: linear  
PCT-US91-01025A-2

Query Match 95.1%; Score 1961; DB 5; Length 527;  
Best Local Similarity 99.7%; Pred. No. 9.3e-186;  
Matches 355; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 21 AASEGNSDCYFCGNSAVRGTHSLTESGASCLPWNMILIGKVTYTAQNPSAALGLGKHNY 80  
Db 172 ACSEGNDCYFCGNSAVRGTHSLTESGASCLPWNMILIGKVTYTAQNPSAALGLGKHNY 231  
Qy 81 CRNPDGDAKPNCHVLKNRRLTWECYCDVPCSTCGLRQYSDQFRIKGLFADIAHPWQA 140  
Db 232 CRNPDGDAKPNCHVLKNRRLTWECYCDVPCSTCGLRQYSDQFRIKGLFADIAHPWQA 291  
Qy 141 AIFAKHRRSPGERFLCGGILISSCWILSAACHFCQERPPPHLTVILGRTYRVVPGEEQK 200  
Db 292 AIFAKHRRSPGERFLCGGILISSCWILSAACHFCQERPPPHLTVILGRTYRVVPGEEQK 351  
Qy 201 FEVEKYIVHKEFDDDDTYDNDIALQLKSDSSRCAQESSVVRTVCLPPADLQLPDWTCECL 260

Db 352 FEVEKIYVHKFEFDDTDYNDIALQLKSDSRCAQESSVVRTVCLPPADLQLPDWTCECL 411  
Qy 261 SGYGKEALSPFYSERLKEAHLVLPSSRCTSQHLLNRTVTDNMLCAGDTRSGGPOANLH 320  
Db 412 SGYGKEALSPFYSERLKEAHLVLPSSRCTSQHLLNRTVTDNMLCAGDTRSGGPOANLH 471  
Qy 321 DACQDGGPLVCLNDGRMTLVGIISWGLGCGQKDVPGVYTKVTNYLDWIRDNRMP 376  
Db 472 DACQDGGPLVCLNDGRMTLVGIISWGLGCGQKDVPGVYTKVTNYLDWIRDNRMP 527

## RESULT 4

5185259-8  
; Patent No. 5185259  
; APPLICANT: GOEDDEL, DAVID V.; KOHR, WILLIAM J.; PENNICA, DIANE;  
; VEIAR, GORDON A.  
; TITLE OF INVENTION: TRUNCATED HUMAN TISSUE PLASMINOGEN  
; ACTIVATOR  
; NUMBER OF SEQUENCES: 15  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/489,855  
; FILING DATE: 02-MAR-1990  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 12,694  
; FILING DATE: 09-FEB-1987  
; APPLICATION NUMBER: 483,052  
; FILING DATE: 07-APR-1983  
; APPLICATION NUMBER: 398,003  
; FILING DATE: 14-JUL-1982  
; APPLICATION NUMBER: 374,860  
; FILING DATE: 05-MAY-1982  
; SEQ ID NO:8  
; LENGTH: 527  
5185259-8

Query Match 95.1%; Score 1961; DB 6; Length 527;  
Best Local Similarity 99.7%; Pred. No. 9.3e-186;  
Matches 355; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 21 AASEGNSDCYFGNGSAVRGTHSLTESGASCLPWNMILIGKVYTAQNPSAQLGLGHNY 80  
Db 172 ACSEGNSDCYFGNGSAVRGTHSLTESGASCLPWNMILIGKVYTAQNPSAQLGLGHNY 231  
Qy 81 CRNPDGDAKWCWHLKNRRLTWECYDVPSCSTCGLRQYSPQFRIKGLFADIASHPWQA 140  
Db 232 CRNPDGDAKWCWHLKNRRLTWECYDVPSCSTCGLRQYSPQFRIKGLFADIASHPWQA 291  
Qy 141 AIFAKHRRSPGERFLCGGILISSCWILSAACHCFQERPPPHLTVILGRTYRVVPEBEQK 200  
Db 292 AIFAKHRRSPGERFLCGGILISSCWILSAACHCFQERPPPHLTVILGRTYRVVPEBEQK 351  
Qy 201 FEVEKIYVHKFEFDDTDYNDIALQLKSDSRCAQESSVVRTVCLPPADLQLPDWTCECL 260  
Db 352 FEVEKIYVHKFEFDDTDYNDIALQLKSDSRCAQESSVVRTVCLPPADLQLPDWTCECL 411  
Qy 261 SGYGKEALSPFYSERLKEAHLVLPSSRCTSQHLLNRTVTDNMLCAGDTRSGGPOANLH 320  
Db 412 SGYGKEALSPFYSERLKEAHLVLPSSRCTSQHLLNRTVTDNMLCAGDTRSGGPOANLH 471  
Qy 321 DACQDGGPLVCLNDGRMTLVGIISWGLGCGQKDVPGVYTKVTNYLDWIRDNRMP 376  
Db 472 DACQDGGPLVCLNDGRMTLVGIISWGLGCGQKDVPGVYTKVTNYLDWIRDNRMP 527

## RESULT 5

US-08-811-949-43  
; Sequence 43, Application US/08811949  
; Patent No. 5840533  
; GENERAL INFORMATION:  
; APPLICANT: NIWA, MINEO  
; APPLICANT: SAITO, YOSHIMASA  
; APPLICANT: SASAKI, HITOSHI  
; APPLICANT: HAYASHI, MASAKO

; APPLICANT: NOTANI, JOUJI  
; APPLICANT: KOBAYASHI, MASAKAZU  
; TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR  
; NUMBER OF SEQUENCES: 67  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
; ADDRESSEE: P.C.  
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400  
; CITY: ARLINGTON  
; STATE: VA  
; COUNTRY: USA  
; ZIP: 22202  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA: US/08/811,949  
; APPLICATION NUMBER: US/08/811,949  
; FILING DATE: 05-MAR-1997  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: OBLON, NORMAN F.  
; REGISTRATION NUMBER: 24,618  
; REFERENCE/DOCKET NUMBER: 18-966-0  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 703-413-3000  
; TELEFAX: 703-413-2220  
; INFORMATION FOR SEQ ID NO: 43:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 562 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-811-949-43

Query Match 95.1%; Score 1961; DB 2; Length 562;  
Best Local Similarity 99.7%; Pred. No. 1e-185;  
Matches 355; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 21 AASEGNSDCYFGNGSAVRGTHSLTESGASCLPWNMILIGKVYTAQNPSAQLGLGHNY 80  
Db 207 ACSEGNSDCYFGNGSAVRGTHSLTESGASCLPWNMILIGKVYTAQNPSAQLGLGHNY 266  
Qy 81 CRNPDGDAKWCWHLKNRRLTWECYDVPSCSTCGLRQYSPQFRIKGLFADIASHPWQA 140  
Db 267 CRNPDGDAKWCWHLKNRRLTWECYDVPSCSTCGLRQYSPQFRIKGLFADIASHPWQA 326  
Qy 141 AIFAKHRRSPGERFLCGGILISSCWILSAACHCFQERPPPHLTVILGRTYRVVPEBEQK 200  
Db 327 AIFAKHRRSPGERFLCGGILISSCWILSAACHCFQERPPPHLTVILGRTYRVVPEBEQK 386  
Qy 201 FEVEKIYVHKFEFDDTDYNDIALQLKSDSRCAQESSVVRTVCLPPADLQLPDWTCECL 260  
Db 387 FEVEKIYVHKFEFDDTDYNDIALQLKSDSRCAQESSVVRTVCLPPADLQLPDWTCECL 446  
Qy 261 SGYGKEALSPFYSERLKEAHLVLPSSRCTSQHLLNRTVTDNMLCAGDTRSGGPOANLH 320  
Db 447 SGYGKEALSPFYSERLKEAHLVLPSSRCTSQHLLNRTVTDNMLCAGDTRSGGPOANLH 506  
Qy 321 DACQDGGPLVCLNDGRMTLVGIISWGLGCGQKDVPGVYTKVTNYLDWIRDNRMP 376  
Db 507 DACQDGGPLVCLNDGRMTLVGIISWGLGCGQKDVPGVYTKVTNYLDWIRDNRMP 562

## RESULT 6

US-08-560-098A-50  
; Sequence 50, Application US/08560098A  
; Patent No. 5976841  
; GENERAL INFORMATION:  
; APPLICANT: WENNDT, Stephan  
; APPLICANT: HEINZEL-WIELAND, Regina  
; APPLICANT: STEPPENS, Gerd Josef

;; TITLE OF INVENTION: Proteins having Fibrinolytic and  
;; TITLE OF INVENTION: Coagulation-inhibiting Properties  
;; NUMBER OF SEQUENCES: 60  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Evenson, McKown, Edwards & Lenahan  
;; STREET: 1200 G Street, N.W., Suite 700  
;; CITY: Washington  
;; STATE: DC  
;; COUNTRY: USA  
;; ZIP: 20005  
;;  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/560,098A  
;; FILING DATE: 17-NOV-1995  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: P 44 40 892.7  
;; FILING DATE: 17-NOV-1994  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: EVANS, Joseph D.  
;; REGISTRATION NUMBER: 26,269  
;; REFERENCE/DOCKET NUMBER: 148/42448  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (202) 628-8800  
;; TELEFAX: (202) 628-8844  
;; INFORMATION FOR SEQ ID NO: 50:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 562 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
;;  
;; US-08-560-098A-50

Query Match 95.1%; Score 1961; DB 2; Length 562;  
Best Local Similarity 99.7%; Pred. No. 1e-185;  
Matches 355; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
Qy 21 AASEGSDCYFGNGSAVYRTHSLTESGASCLPWNMILIGKVTYTAQNPQAALGLGKHNY 80  
Db 207 ACSEGSDCYFGNGSAVYRTHSLTESGASCLPWNMILIGKVTYTAQNPQAALGLGKHNY 266  
  
Qy 81 CRNPDGAKPWCHLVKNRRLTWECYDVPSCSTGLRQYSDQFRIKGLFADIAHPWQA 140  
Db 267 CRNPDGAKPWCHLVKNRRLTWECYDVPSCSTGLRQYSDQFRIKGLFADIAHPWQA 326  
  
Qy 141 AIFAKHRRSPGERFLCGGILISSCWILSAACHFCQERPPPHLTVILGRTYRVVPGEEQK 200  
Db 327 AIFAKHRRSPGERFLCGGILISSCWILSAACHFCQERPPPHLTVILGRTYRVVPGEEQK 386  
  
Qy 201 FEVEKYIVHKEFDDTDYNDIALQLKSDSRCAQESSVVRTVCLPPADLQLPDWTCECL 260  
Db 387 FEVEKYIVHKEFDDTDYNDIALQLKSDSRCAQESSVVRTVCLPPADLQLPDWTCECL 446  
  
Qy 261 SGYKHEALSPFYSERLKEAHLVLPSSRCTSOHLNRTVTDNMLCAGDTRSGGPQANLH 320  
Db 447 SGYKHEALSPFYSERLKEAHLVLPSSRCTSOHLNRTVTDNMLCAGDTRSGGPQANLH 506  
  
Qy 321 DACQDGGPLVCLNDGRMTLVGIIISWGLGCGKQDVPVYTKVTNYLDWIRDNRNP 376  
Db 507 DACQDGGPLVCLNDGRMTLVGIIISWGLGCGKQDVPVYTKVTNYLDWIRDNRNP 562

RESULT 7  
US-08-883-795A-38  
; Sequence 38, Application US/08883795A  
; Patent No. 5985607  
; GENERAL INFORMATION:  
; APPLICANT: Delcuve, Genevieve  
; ATTORNEY: Awang, Gregor

;; TITLE OF INVENTION: Recombinant DNA Molecules and Expression  
;; TITLE OF INVENTION: Vectors for Tissue Plasminogen Activator  
;; NUMBER OF SEQUENCES: 39  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: BERESKIN & PARR  
;; STREET: 40 King Street West  
;; CITY: Toronto  
;; STATE: Ontario  
;; COUNTRY: Canada  
;; ZIP: M5H 3Y2  
;;  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.25  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/883,795A  
;; FILING DATE: 27-JUN-1997  
;; CLASSIFICATION: 435  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Gravelle, Micheline  
;; REGISTRATION NUMBER: 40,261  
;; REFERENCE/DOCKET NUMBER: 7841-062  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (416) 364-7311  
;; TELEFAX: (416) 361-1398  
;; INFORMATION FOR SEQ ID NO: 38:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 562 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein (tpa)  
;;  
;; US-08-883-795A-38

Query Match 95.1%; Score 1961; DB 2; Length 562;  
Best Local Similarity 99.7%; Pred. No. 1e-185;  
Matches 355; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
Qy 21 AASEGSDCYFGNGSAVYRTHSLTESGASCLPWNMILIGKVTYTAQNPQAALGLGKHNY 80  
Db 207 ACSEGSDCYFGNGSAVYRTHSLTESGASCLPWNMILIGKVTYTAQNPQAALGLGKHNY 266  
  
Qy 81 CRNPDGAKPWCHLVKNRRLTWECYDVPSCSTGLRQYSDQFRIKGLFADIAHPWQA 140  
Db 267 CRNPDGAKPWCHLVKNRRLTWECYDVPSCSTGLRQYSDQFRIKGLFADIAHPWQA 326  
  
Qy 141 AIFAKHRRSPGERFLCGGILISSCWILSAACHFCQERPPPHLTVILGRTYRVVPGEEQK 200  
Db 327 AIFAKHRRSPGERFLCGGILISSCWILSAACHFCQERPPPHLTVILGRTYRVVPGEEQK 386  
  
Qy 201 FEVEKYIVHKEFDDTDYNDIALQLKSDSRCAQESSVVRTVCLPPADLQLPDWTCECL 260  
Db 387 FEVEKYIVHKEFDDTDYNDIALQLKSDSRCAQESSVVRTVCLPPADLQLPDWTCECL 446  
  
Qy 261 SGYKHEALSPFYSERLKEAHLVLPSSRCTSOHLNRTVTDNMLCAGDTRSGGPQANLH 320  
Db 447 SGYKHEALSPFYSERLKEAHLVLPSSRCTSOHLNRTVTDNMLCAGDTRSGGPQANLH 506  
  
Qy 321 DACQDGGPLVCLNDGRMTLVGIIISWGLGCGKQDVPVYTKVTNYLDWIRDNRNP 376  
Db 507 DACQDGGPLVCLNDGRMTLVGIIISWGLGCGKQDVPVYTKVTNYLDWIRDNRNP 562

RESULT 8  
5185259-3  
; Patent No. 5185259  
; APPLICANT: GOEDEL, DAVID V.; KOHR, WILLIAM J.; PENNICA, DIANE;  
; VEHAR, GORDON A.  
; TITLE OF INVENTION: TRUNCATED HUMAN TISSUE PLASMINOGEN  
; ACTIVATOR  
; NUMBER OF SEQUENCES: 15  
; CURRENT APPLICATION DATA:





; FILING DATE: 06-JUN-1988

; SEQ ID NO: 5;

; LENGTH: 562

5244676-5

Query Match 95.0%; Score 1958; DB 6; Length 562;  
Best Local Similarity 99.4%; Pred. No. 2e-185;  
Matches 354; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 21 AASEGNSCYFGNGSAVYRGTHSLTESGASCLPWNMILIGKVTYTAQNPQAALGLGKHYN 80  
Db 207 SCSEGNSCYFGNGSAVYRGTHSLTESGASCLPWNMILIGKVTYTAQNPQAALGLGKHYN 266  
Qy 81 CRNPDGDAKPVCHLVKNRRLTWYCDVPSCTGLRQYSDQFRIKGGFLFADIASHPWQA 140  
Db 267 CRNPDGDAKPVCHLVKNRRLTWYCDVPSCTGLRQYSDQFRIKGGFLFADIASHPWQA 326  
Qy 141 AIFAKHRSPPGERFLCGGILISSCWILSAAHCFQERFPPHLLTVILGRTYRVVPGEEQK 200  
Db 327 AIFAKHRSPPGERFLCGGILISSCWILSAAHCFQERFPPHLLTVILGRTYRVVPGEEQK 386  
Qy 201 FEVEKIIVHKEFDDTDYNDIALQLKSDSRCAQESSVVRTVCLPPADLQLPDWTCECEL 260  
Db 387 FEVEKIIVHKEFDDTDYNDIALQLKSDSRCAQESSVVRTVCLPPADLQLPDWTCECEL 446  
Qy 261 SGYGKEALSPFYSERLKEAHVRLYPSSRCTSOHLNRTVTDNMLCAGDTRSGGPOANLH 320  
Db 447 SGYGKEALSPFYSERLKEAHVRLYPSSRCTSOHLNRTVTDNMLCAGDTRSGGPOANLH 506  
Qy 321 DACQDSGGPLVCLNDGRMTLVGIISWGLGCGQKDVPGVYTKVTNYLDWIRDNMRP 376  
Db 507 DACQDSGGPLVCLNDGRMTLVGIISWGLGCGQKDVPGVYTKVTNYLDWIRDNMRP 562

## RESULT 12

US-08-811-949-45  
; Sequence 45, Application US/08811949  
; Patent No. 5840533  
; GENERAL INFORMATION:  
; APPLICANT: NIWA, MINEO  
; APPLICANT: SAITO, YOSHIMASA  
; APPLICANT: SASAKI, HITOSHI  
; APPLICANT: HAYASHI, MASAKO  
; APPLICANT: NOTANI, JOUJI  
; APPLICANT: KOBAYASHI, MASAKAZU  
; TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR  
; NUMBER OF SEQUENCES: 67  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: P.C.  
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400  
; CITY: ARLINGTON  
; STATE: VA  
; COUNTRY: USA  
; ZIP: 22202  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/811,949  
; FILING DATE: 05-MAR-1997  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: OBLON, NORMAN F.  
; REGISTRATION NUMBER: 24,618  
; REFERENCE/DOCKET NUMBER: 18-966-0  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 703-413-3000  
; TELEFAX: 703-413-2220  
; INFORMATION FOR SEQ ID NO: 45:  
; SEQUENCE CHARACTERISTICS:

LENGTH: 355 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-811-949-45

Query Match 95.0%; Score 1957; DB 2; Length 355;  
Best Local Similarity 100.0%; Pred. No. 1.3e-185;  
Matches 354; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 23 SEGNSDCYFGNGSAVYRGTHSLTESGASCLPWNMILIGKVTYTAQNPQAALGLGKHYNCR 82  
Db 2 SEGNSDCYFGNGSAVYRGTHSLTESGASCLPWNMILIGKVTYTAQNPQAALGLGKHYNCR 61  
Qy 83 NPQGDGDAKPVCHLVKNRRLTWYCDVPSCTGLRQYSDQFRIKGGFLFADIASHPWQA 142  
Db 62 NPQGDGDAKPVCHLVKNRRLTWYCDVPSCTGLRQYSDQFRIKGGFLFADIASHPWQA 121  
Qy 143 FAKHRSPPGERFLCGGILISSCWILSAAHCFQERFPPHLLTVILGRTYRVVPGEEQKFE 202  
Db 122 FAKHRSPPGERFLCGGILISSCWILSAAHCFQERFPPHLLTVILGRTYRVVPGEEQKFE 181  
Qy 203 VEKIIVHKEFDDTDYNDIALQLKSDSRCAQESSVVRTVCLPPADLQLPDWTCELSG 262  
Db 182 VEKIIVHKEFDDTDYNDIALQLKSDSRCAQESSVVRTVCLPPADLQLPDWTCELSG 241  
Qy 263 YGKEALSPFYSERLKEAHVRLYPSSRCTSOHLNRTVTDNMLCAGDTRSGGPOANLHDA 322  
Db 242 YGKEALSPFYSERLKEAHVRLYPSSRCTSOHLNRTVTDNMLCAGDTRSGGPOANLHDA 301  
Qy 323 CQDSGGPLVCLNDGRMTLVGIISWGLGCGQKDVPGVYTKVTNYLDWIRDNMRP 376  
Db 302 CQDSGGPLVCLNDGRMTLVGIISWGLGCGQKDVPGVYTKVTNYLDWIRDNMRP 355

## RESULT 13

US-08-811-949-49  
; Sequence 49, Application US/08811949  
; Patent No. 5840533  
; GENERAL INFORMATION:  
; APPLICANT: NIWA, MINEO  
; APPLICANT: SAITO, YOSHIMASA  
; APPLICANT: SASAKI, HITOSHI  
; APPLICANT: HAYASHI, MASAKO  
; APPLICANT: NOTANI, JOUJI  
; APPLICANT: KOBAYASHI, MASAKAZU  
; TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR  
; NUMBER OF SEQUENCES: 67  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: P.C.  
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400  
; CITY: ARLINGTON  
; STATE: VA  
; COUNTRY: USA  
; ZIP: 22202  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/811,949  
; FILING DATE: 05-MAR-1997  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: OBLON, NORMAN F.  
; REGISTRATION NUMBER: 24,618  
; REFERENCE/DOCKET NUMBER: 18-966-0  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 703-413-3000  
; TELEFAX: 703-413-2220  
; INFORMATION FOR SEQ ID NO: 49:

```
; SEQUENCE CHARACTERISTICS:
; LENGTH: 437 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-811-949-49

Query Match          95.0%; Score 1957; DB 2; Length 437;
Best Local Similarity 99.4%; Pred. No. 1.8e-185;
Matches 354; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 21 AASEGNSDCYFGNSAYRGTHSLTESGASCLPWNMILIGKVTYTAQNPQAALGLGKHY 80
Db 82 ACSEGNSDCYFGNSAYRGTHSLTESGASCLPWNMILIGKVTYTAQNPQAALGLGKHY 141
Qy 81 CRNPDGDAKWPCHVKNRRLTWECYCDVPSCTGRLQVYQFRIKGLFADIAHPWQA 140
Db 142 CRNPDGDAKWPCHVKNRRLTWECYCDVPSCTGRLQVYQFRIKGLFADIAHPWQA 201
Qy 141 AIFAKHRRSPGERFLCGGILISSCWILSSAHCFQERPPPHLTIVILGRTYRVVPEEEOK 200
Db 202 AIFAKHRRSPGERFLCGGILISSCWILSSAHCFQERPPPHLTIVILGRTYRVVPEEEOK 261
Qy 201 FEVEKYIVHKEFDDTDYNDIALIOLKSDSRCAQESSVVRTVCLPPADLQLPDWTCEL 260
Db 262 FEVEKYIVHKEFDDTDYNDIALIOLKSDSRCAQESSVVRTVCLPPADLQLPDWTCEL 321
Qy 261 SGYKHEALSPFYSERLKEAHVRLYPSRCTSOHLNRTVTDNMLCAGDTRSGGPOANLH 320
Db 322 SGYKHEALSPFYSERLKEAHVRLYPSRCTSOHLNRTVTDNMLCAGDTRSGGPOANLH 381
Qy 321 DACQDGGPLVCLNDGRMTLVGIISWGLCGQKDVFGVYTKVTNYLDWIRDNRMP 376
Db 382 DACQDGGPLVCLNDGRMTLVGIISWGLCGQKDVFGVYTKVTNYLDWIRDNRMP 437

RESULT 14
US-08-811-949-65
; Sequence 65, Application US/08811949
; Patent No. 5840533
; GENERAL INFORMATION:
; APPLICANT: NIWA, MINEO
; APPLICANT: SAITO, YOSHIMASA
; APPLICANT: SASAKI, HITOSHI
; APPLICANT: HAYASHI, MASAKO
; APPLICANT: NOTANI, JOUJI
; APPLICANT: KOBAYASHI, MASAKAZU
; TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR
; NUMBER OF SEQUENCES: 67
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: P.C.
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/811,949
; FILING DATE: 05-MAR-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 18-966-0
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220

; INFORMATION FOR SEQ ID NO: 65:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 389 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-811-949-65

Query Match          94.9%; Score 1955.5; DB 2; Length 389;
Best Local Similarity 93.2%; Pred. No. 2.1e-185;
Matches 359; Conservative 4; Mismatches 11; Indels 11; Gaps 2;

Qy 3 KTAIAIAVALAG--FATVAQ-----AASEGNSDCYFGNSAYRGTHSLTESGASCL 51
Db 5 KRGLCCVLLLCGALFVSPSQEIHARFRRGARSEGNSDCYFGNSAYRGTHSLTESGASCL 64
Qy 52 PWNMILIGKVTYTAQNPQAALGLGKHYCRNPDGDAKWPCHVKNRRLTWECYCDVPSCS 111
Db 65 PWNMILIGKVTYTAQNPQAALGLGKHYCRNPDGDAKWPCHVKNRRLTWECYCDVPSCS 124
Qy 112 TCGLROYSQFRIKGLFADIAHPWQAIFAKHRRSPGERFLCGGILISSCWILSSAH 171
Db 125 TCGLROYSQFRIKGLFADIAHPWQAIFAKHRRSPGERFLCGGILISSCWILSSAH 184
Qy 172 CPQERPPPHLTIVILGRTYRVVPEEEOKEVEKYIVHKEFDDTDYNDIALIOLKSDSS 231
Db 185 CPQERPPPHLTIVILGRTYRVVPEEEOKEVEKYIVHKEFDDTDYNDIALIOLKSDSS 244
Qy 232 RQAQESSVVRTVCLPPADLQLPDWTCELSGYKHEALSPFYSERLKEAHVRLYPSRCT 291
Db 245 RQAQESSVVRTVCLPPADLQLPDWTCELSGYKHEALSPFYSERLKEAHVRLYPSRCT 304
Qy 292 SQHLNRTVTDNMLCAGDTRSGGPOANLHDAQCGDGGPLVCLNDGRMTLVGIISWGLGC 351
Db 305 SQHLNRTVTDNMLCAGDTRSGGPOANLHDAQCGDGGPLVCLNDGRMTLVGIISWGLGC 364
Qy 352 GQKDVPGVYTKVTNYLDWIRDNRMP 376
Db 365 GQKDVPGVYTKVTNYLDWIRDNRMP 389

RESULT 15
US-08-811-949-55
; Sequence 55, Application US/08811949
; Patent No. 5840533
; GENERAL INFORMATION:
; APPLICANT: NIWA, MINEO
; APPLICANT: SAITO, YOSHIMASA
; APPLICANT: SASAKI, HITOSHI
; APPLICANT: HAYASHI, MASAKO
; APPLICANT: NOTANI, JOUJI
; APPLICANT: KOBAYASHI, MASAKAZU
; TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR
; NUMBER OF SEQUENCES: 67
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: P.C.
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/811,949
; FILING DATE: 05-MAR-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F.
```

REGISTRATION NUMBER: 24,618  
REFERENCE/DOCKET NUMBER: 18-966-0  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-413-3000  
TELEFAX: 703-413-2220  
INFORMATION FOR SEQ ID NO: 55:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 437 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-811-949-55

Query Match 94.8%; Score 1954; DB 2; Length 437;  
Best Local Similarity 99.4%; Pred. No. 3.5e-185;  
Matches 354; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy	21	AASEGNSDCYFGNGSAYRGTHSLTESGASCLPWNMSMILIGKVYTAQNPSAQALGLGKHNY	80
Db	82	ACSEGNSDCYFGNGSAYRGTHSLTESGASCLPWNMSMILIGKVYTAQNPSAQALGLGKHNY	141
Qy	81	CRNPDGDAKPWCHVLKNRRLTWECYCDVPSCSTCGLRQYSQPQFRIKGLFADIASHPWQA	140
Db	142	CRNPDGDAKPWCHVLKNRRLTWECYCDVPSCSTCGLRQYSQPQFRIKGLFADIASHPWQA	201
Qy	141	AIFAKHRRSPGERFLCGGILISSCWILSAAHCFQERPPPHLTVILGRTYRVVPGEEOK	200
Db	202	AIFAKHRRSPGERFLCGGILISSCWILSAAHCFQERPPPHLTVILGRTYRVVPGEEOK	261
Qy	201	FEVEKYIVHKEFDDDTYDNDIALQLKSDSRCAQESSVVRTVCLPPADLQLPDWTCEEL	260
Db	262	FEVEKYIVHKEFDDDTYDNDIALQLKSDSRCAQESSVVRTVCLPPADLQLPDWTCEEL	321
Qy	261	SGYKGHEALSPFYSERLKEAHVRLYPSSRCTSOHLNRTVTDNMLCAGDTRSGGPQANLH	320
Db	322	SGYKGHEALSPFYSERLKEAHVRLYPSSRCTSOHLNRTVTDNMLCAGDTRSGGPQANLH	381
Qy	321	DACOGDSGGPLVCLNDGRMTLVGLISWGLGCGKDXDVGVTYKVTNYLDWIRDNNRP	376
Db	382	DACOGDSGGPLVCLNDGRMTLVGLISWGLGCGKDXDVGVTYKVTNYLDWIRDNNRP	437

Search completed: August 8, 2003, 17:15:51  
Job time : 21 secs

Result No.	Score	Query		Length	DB	ID	Description
		Match	%				
1	2061	100.0		377	11	US-09-987-455-8	Sequence 8, Appli
2	1961	95.1		372	9	US-09-084-491A-3	Sequence 3, Appli
3	1961	95.1		372	14	US-10-102-704-3	Sequence 3, Appli
4	1961	95.1		527	11	US-09-987-457-18	Sequence 18, Appli
5	1961	95.1		527	11	US-09-987-455-19	Sequence 19, Appli
6	1961	95.1		562	9	US-09-969-271-7	Sequence 7, Appli
7	1981	95.1		562	15	US-10-193-656-8	Sequence 8, Appli
8	1957	95.0		354	11	US-09-987-457-10	Sequence 10, Appli
9	1957	95.0		354	11	US-09-987-455-11	Sequence 11, Appli
10	1952	94.7		562	10	US-09-974-298-14	Sequence 14, Appli
11	1875.5	91.0		343	11	US-09-987-457-14	Sequence 14, Appli
12	1875.5	91.0		343	11	US-09-987-455-15	Sequence 15, Appli
13	1865.5	90.5		343	11	US-09-987-457-15	Sequence 15, Appli
14	1865.5	90.5		343	11	US-09-987-455-16	Sequence 16, Appli
15	1850.5	89.8		339	11	US-09-987-457-12	Sequence 12, Appli



## RESULT 4

US-09-987-457-18  
; Sequence 18, Application US/09987457  
; Publication No. US20030013150A1  
; GENERAL INFORMATION:  
; APPLICANT: Manosroi, Aranya  
; APPLICANT: Manosroi, Jiradej  
; APPLICANT: Tayapiwatana, Chatchai  
; APPLICANT: Goetz, Friedrich  
; APPLICANT: Werner, Rolf-Guenther  
; TITLE OF INVENTION: Methods for Large Scale Protein Production in Prokaryotes  
; FILE REFERENCE: 0652.2180001  
; CURRENT APPLICATION NUMBER: US/09/987,457  
; CURRENT FILING DATE: 2001-11-14  
; PRIOR FILING DATE: 2001-11-14  
; PRIOR FILING DATE: 2001-02-15  
; PRIOR APPLICATION NUMBER: GB 00 27 782.2  
; PRIOR FILING DATE: 2000-11-14  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 18  
; LENGTH: 527  
; TYPE: PRT  
; ORGANISM: Homo sapiens (tPA)  
US-09-987-457-18.

Query Match 95.1%; Score 1961; DB 11; Length 527;  
Best Local Similarity 99.7%; Pred. No. 7.9e-183;  
Matches 355; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
Qy 21 AASEGNSDCYFGNGSAYRGTHSLTESGASCLPWNMSMLIGKVYTAQNPSAQLGLGKHNY 80  
Db 172 ACSEGNSDCYFGNGSAYRGTHSLTESGASCLPWNMSMLIGKVYTAQNPSAQLGLGKHNY 231  
  
Qy 81 CRNPDGAKPWCHVKNRRLTWECYDVPSCSTCGLROYSQOPRIKGGLPADIASHPWQA 140  
Db 232 CRNPDGAKPWCHVKNRRLTWECYDVPSCSTCGLROYSQOPRIKGGLPADIASHPWQA 291  
  
Qy 141 AIFAKHRRSPGERFLCGGILISSCWILSAACHFOERPPPHLTVILGRYRVVPGEEQK 200  
Db 292 AIFAKHRRSPGERFLCGGILISSCWILSAACHFOERPPPHLTVILGRYRVVPGEEQK 351  
  
Qy 201 FEVEKYIVHKEFDYDNDIALQLKSDSSRCAQESSVVRTVCLPPADLQLPDWTCECL 260  
Db 352 FEVEKYIVHKEFDYDNDIALQLKSDSSRCAQESSVVRTVCLPPADLQLPDWTCECL 411  
  
Qy 261 SGYKKEALSPFYSERLKEAHVRLYPSRCTSQHLLNRTVTDNMLCAGDTRSGGPOANLH 320  
Db 412 SGYKKEALSPFYSERLKEAHVRLYPSRCTSQHLLNRTVTDNMLCAGDTRSGGPOANLH 471  
  
Qy 321 DACQDGGPLVCLNDGRMTLVGIISWGLGCGQKDVPGVTKVTNYLDWIRDNRMP 376  
Db 472 DACQDGGPLVCLNDGRMTLVGIISWGLGCGQKDVPGVTKVTNYLDWIRDNRMP 527

## RESULT 5

US-09-987-455-19  
; Sequence 19, Application US/09987455  
; Publication No. US20030049729A1  
; GENERAL INFORMATION:  
; APPLICANT: Aranya Manosroi  
; APPLICANT: Jiradej Manosroi  
; APPLICANT: Chatchai Tayapiwatana  
; APPLICANT: Friedrich Goetz  
; APPLICANT: Rolf-Guenther Werner  
; TITLE OF INVENTION: Methods for Large Scale Production of Recombinant  
; FILE REFERENCE: 0652.2190001  
; CURRENT APPLICATION NUMBER: US/09/987,455  
; CURRENT FILING DATE: 2001-11-14  
; PRIOR FILING DATE: 2001-02-15

; PRIOR APPLICATION NUMBER: GB 0027779.8  
; PRIOR FILING DATE: 2000-11-14  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 19  
; LENGTH: 527  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-987-455-19

Query Match 95.1%; Score 1961; DB 11; Length 527;  
Best Local Similarity 99.7%; Pred. No. 7.9e-183;  
Matches 355; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
Qy 21 AASEGNSDCYFGNGSAYRGTHSLTESGASCLPWNMSMLIGKVYTAQNPSAQLGLGKHNY 80  
Db 172 ACSEGNSDCYFGNGSAYRGTHSLTESGASCLPWNMSMLIGKVYTAQNPSAQLGLGKHNY 231  
  
Qy 81 CRNPDGAKPWCHVKNRRLTWECYDVPSCSTCGLROYSQOPRIKGGLPADIASHPWQA 140  
Db 232 CRNPDGAKPWCHVKNRRLTWECYDVPSCSTCGLROYSQOPRIKGGLPADIASHPWQA 291  
  
Qy 141 AIFAKHRRSPGERFLCGGILISSCWILSAACHFOERPPPHLTVILGRYRVVPGEEQK 200  
Db 292 AIFAKHRRSPGERFLCGGILISSCWILSAACHFOERPPPHLTVILGRYRVVPGEEQK 351  
  
Qy 201 FEVEKYIVHKEFDYDNDIALQLKSDSSRCAQESSVVRTVCLPPADLQLPDWTCECL 260  
Db 352 FEVEKYIVHKEFDYDNDIALQLKSDSSRCAQESSVVRTVCLPPADLQLPDWTCECL 411  
  
Qy 261 SGYKKEALSPFYSERLKEAHVRLYPSRCTSQHLLNRTVTDNMLCAGDTRSGGPOANLH 320  
Db 412 SGYKKEALSPFYSERLKEAHVRLYPSRCTSQHLLNRTVTDNMLCAGDTRSGGPOANLH 471  
  
Qy 321 DACQDGGPLVCLNDGRMTLVGIISWGLGCGQKDVPGVTKVTNYLDWIRDNRMP 376  
Db 472 DACQDGGPLVCLNDGRMTLVGIISWGLGCGQKDVPGVTKVTNYLDWIRDNRMP 527

## RESULT 6

US-09-969-271-7  
; Sequence 7, Application US/09969271  
; Patent No. US20020098179A1  
; GENERAL INFORMATION:  
; APPLICANT: Pfizer Inc. (All designated States except GB and EP (GB));  
; APPLICANT: Pfizer Limited (GB and EP (GB) only)  
; TITLE OF INVENTION: Pharmaceutical Combinations  
; FILE REFERENCE: PCS10951APME  
; CURRENT APPLICATION NUMBER: US/09/969,271  
; CURRENT FILING DATE: 2001-10-01  
; PRIOR FILING DATE: 2000-10-17  
; NUMBER OF SEQ ID NOS: 7  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 7  
; LENGTH: 562  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-969-271-7

Query Match 95.1%; Score 1961; DB 9; Length 562;  
Best Local Similarity 99.7%; Pred. No. 8.6e-183;  
Matches 355; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
Qy 21 AASEGNSDCYFGNGSAYRGTHSLTESGASCLPWNMSMLIGKVYTAQNPSAQLGLGKHNY 80  
Db 207 ACSEGNSDCYFGNGSAYRGTHSLTESGASCLPWNMSMLIGKVYTAQNPSAQLGLGKHNY 266  
  
Qy 81 CRNPDGAKPWCHVKNRRLTWECYDVPSCSTCGLROYSQOPRIKGGLPADIASHPWQA 140  
Db 267 CRNPDGAKPWCHVKNRRLTWECYDVPSCSTCGLROYSQOPRIKGGLPADIASHPWQA 326  
  
Qy 141 AIFAKHRRSPGERFLCGGILISSCWILSAACHFOERPPPHLTVILGRYRVVPGEEQK 200

Db 327 AIFAKHRRSPGERFLCGGILISSCWILSAAHCFQERPPHLLTVILGRTYRVVPEEQK 386  
Qy 201 FEVEKYIVHKEFDDDDTYNDIALLQLKSDSSRCQAQESSVVRTVCLPPADLQLPDWTCECL 260  
Db 387 FEVEKYIVHKEFDDDDTYNDIALLQLKSDSSRCQAQESSVVRTVCLPPADLQLPDWTCECL 446  
Qy 261 SGYKHEALSPFYSERLKEAHVRLYPSSRCTSOHLNRTVTDNMLCAGDTRSGGQANLH 320  
Db 447 SGYKHEALSPFYSERLKEAHVRLYPSSRCTSOHLNRTVTDNMLCAGDTRSGGQANLH 506  
Qy 321 DACQDGGPLVCLNDGRMTLVGIISWGLGCGKDXVPGVYTKVTNYLDWIRDNRMP 376  
Db 507 DACQDGGPLVCLNDGRMTLVGIISWGLGCGKDXVPGVYTKVTNYLDWIRDNRMP 562

RESULT 7  
US-10-193-656-8  
; Sequence 8, Application US/10193656  
; Publication No. US20030096733A1  
; GENERAL INFORMATION:  
; APPLICANT: NY, Tor  
; APPLICANT: HOLMDAHL, Rikard  
; APPLICANT: Li, Jinan  
; TITLE OF INVENTION: NOVEL DRUG TARGETS FOR ARTHRITIS  
; FILE REFERENCE: 3810/1J577-US3  
; CURRENT APPLICATION NUMBER: US/10/193,656  
; CURRENT FILING DATE: 2002-07-10  
; PRIOR APPLICATION NUMBER: US 60/304,461  
; PRIOR FILING DATE: 2001-07-10  
; PRIOR APPLICATION NUMBER: US 60/304,490  
; PRIOR FILING DATE: 2001-07-10  
; PRIOR APPLICATION NUMBER: US 60/305,182  
; PRIOR FILING DATE: 2001-07-13  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: Patent in version 3.1  
; SEQ ID NO 8  
; LENGTH: 562  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; PUBLICATION INFORMATION:  
; DATABASE ACCESSION NUMBER: GenBank / P00750  
; DATABASE ENTRY DATE: 1986-07-21  
; RELEVANT RESIDUES: (1)..(562)

US-10-193-656-8

Query Match 95.1%; Score 1961; DB 15; Length 562;  
Best Local Similarity 99.7%; Pred. No. 8.6e-183;  
Matches 355; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 21 AASEGNSDCYFGNGSAYRGTHSLTESGASCLPWNMILIGKVTYTAQNPASQAALGLGKHNY 80  
Db 207 ACSEGNSDCYFGNGSAYRGTHSLTESGASCLPWNMILIGKVTYTAQNPASQAALGLGKHNY 266  
Qy 81 CRNPDGAKPWCHVLKNRRLTWECYDVPSCSTCGLRQYVSPQFRKGLFADIASHPWQA 140  
Db 267 CRNPDGAKPWCHVLKNRRLTWECYDVPSCSTCGLRQYVSPQFRKGLFADIASHPWQA 326  
Qy 141 AIFAKHRRSPGERFLCGGILISSCWILSAAHCFQERPPHLLTVILGRTYRVVPEEQK 200  
Db 327 AIFAKHRRSPGERFLCGGILISSCWILSAAHCFQERPPHLLTVILGRTYRVVPEEQK 386  
Qy 201 FEVEKYIVHKEFDDDDTYNDIALLQLKSDSSRCQAQESSVVRTVCLPPADLQLPDWTCECL 260  
Db 387 FEVEKYIVHKEFDDDDTYNDIALLQLKSDSSRCQAQESSVVRTVCLPPADLQLPDWTCECL 446  
Qy 261 SGYKHEALSPFYSERLKEAHVRLYPSSRCTSOHLNRTVTDNMLCAGDTRSGGQANLH 320  
Db 447 SGYKHEALSPFYSERLKEAHVRLYPSSRCTSOHLNRTVTDNMLCAGDTRSGGQANLH 506  
Qy 321 DACQDGGPLVCLNDGRMTLVGIISWGLGCGKDXVPGVYTKVTNYLDWIRDNRMP 376  
Db 507 DACQDGGPLVCLNDGRMTLVGIISWGLGCGKDXVPGVYTKVTNYLDWIRDNRMP 562

RESULT 8  
US-09-987-457-10  
; Sequence 10, Application US/09987457  
; Publication No. US20030013150A1  
; GENERAL INFORMATION:  
; APPLICANT: Manosroi, Aranya  
; APPLICANT: Manosroi, Jiradej  
; APPLICANT: Tayapiwatana, Chatchai  
; APPLICANT: Goetz, Friedrich  
; APPLICANT: Werner, Rolf-Guenther  
; TITLE OF INVENTION: Methods for Large Scale Protein Production in Prokaryotes  
; FILE REFERENCE: 0652.2180001  
; CURRENT APPLICATION NUMBER: US/09/987,457  
; CURRENT FILING DATE: 2001-11-14  
; PRIOR APPLICATION NUMBER: 60/268,573  
; PRIOR FILING DATE: 2001-02-15  
; PRIOR APPLICATION NUMBER: GB 00 27 782.2  
; PRIOR FILING DATE: 2000-11-14  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: Patent in Ver. 2.1  
; SEQ ID NO 10  
; LENGTH: 354  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: part of the  
; OTHER INFORMATION: recombinant K2S molecule

Query Match 95.0%; Score 1957; DB 11; Length 354;  
Best Local Similarity 100.0%; Pred. No. 1.1e-182;  
Matches 354; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 23 SEGNSDCYFGNGSAYRGTHSLTESGASCLPWNMILIGKVTYTAQNPASQAALGLGKHNYCR 82  
Db 1 SEGNSDCYFGNGSAYRGTHSLTESGASCLPWNMILIGKVTYTAQNPASQAALGLGKHNYCR 60  
Qy 83 NPDGAKPWCHVLKNRRLTWECYDVPSCSTCGLRQYVSPQFRKGLFADIASHPWQA 142  
Db 61 NPDGAKPWCHVLKNRRLTWECYDVPSCSTCGLRQYVSPQFRKGLFADIASHPWQA 120  
Qy 143 FAKHRRSPGERFLCGGILISSCWILSAAHCFQERPPHLLTVILGRTYRVVPEEQKPE 202  
Db 121 FAKHRRSPGERFLCGGILISSCWILSAAHCFQERPPHLLTVILGRTYRVVPEEQKPE 180  
Qy 203 VEKYIVHKEFDDDDTYNDIALLQLKSDSSRCQAQESSVVRTVCLPPADLQLPDWTCECLSG 262  
Db 181 VEKYIVHKEFDDDDTYNDIALLQLKSDSSRCQAQESSVVRTVCLPPADLQLPDWTCECLSG 240  
Qy 263 YGKHEALSPFYSERLKEAHVRLYPSSRCTSOHLNRTVTDNMLCAGDTRSGGQANLHDA 322  
Db 241 YGKHEALSPFYSERLKEAHVRLYPSSRCTSOHLNRTVTDNMLCAGDTRSGGQANLHDA 300  
Qy 323 CQDSDGGPLVCLNDGRMTLVGIISWGLGCGKDXVPGVYTKVTNYLDWIRDNRMP 376  
Db 301 CQDSDGGPLVCLNDGRMTLVGIISWGLGCGKDXVPGVYTKVTNYLDWIRDNRMP 354

RESULT 9  
US-09-987-455-11  
; Sequence 11, Application US/09987455  
; Publication No. US20030049729A1  
; GENERAL INFORMATION:  
; APPLICANT: Aranya Manosroi  
; APPLICANT: Jiradej Manosroi  
; APPLICANT: Chatchai Tayapiwatana  
; APPLICANT: Friedrich Goetz  
; APPLICANT: Rolf-Guenther Werner  
; TITLE OF INVENTION: Methods for Large Scale Production of Recombinant  
; TITLE OF INVENTION: DNA-Derived tpa or K2S Molecules  
; FILE REFERENCE: 0652.2190001



```
; CURRENT APPLICATION NUMBER: US/09/987,455
; CURRENT FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: 60/268,574
; PRIOR FILING DATE: 2001-02-15
; PRIOR APPLICATION NUMBER: GB 0027779.8
; PRIOR FILING DATE: 2000-11-14
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 11
; LENGTH: 354
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: K2S 174-527
US-09-987-455-11

Query Match          95.0%; Score 1957; DB 11; Length 354;
Best Local Similarity 100.0%; Pred. No. 1.1e-182;
Matches 354; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 23 SEGNSDCYFGNGSAYRGTHSLTESGASCLPWNMILIGKVYTAQNPSAALGLGKHNYCR 82
Db 1 SEGNSDCYFGNGSAYRGTHSLTESGASCLPWNMILIGKVYTAQNPSAALGLGKHNYCR 60

Qy 83 NPDGDAPKFWCHLVKNRRLTWECYDVPCSTCGLRQYQSQPQRIKGLFADIASHPWQAAI 142
Db 61 NPDGDAPKFWCHLVKNRRLTWECYDVPCSTCGLRQYQSQPQRIKGLFADIASHPWQAAI 120

Qy 143 FAKHRRSPGERFLCGGILISSCWILSAAHCFQERPPPHLTVILGRYRVVPGEEQKEE 202
Db 121 FAKHRRSPGERFLCGGILISSCWILSAAHCFQERPPPHLTVILGRYRVVPGEEQKEE 180

Qy 203 VEKYIVHKEFDDDDTYNDIALQLKSDSSRCAQESSVVRTVCLPPADLPDWTCELSG 262
Db 181 VEKYIVHKEFDDDDTYNDIALQLKSDSSRCAQESSVVRTVCLPPADLPDWTCELSG 240

Qy 263 YGKHEALSPFYSERLKEAHVRLYPSRCTSOHLNRTVTDNMLCAGDTRSGGPQANLHDA 322
Db 241 YGKHEALSPFYSERLKEAHVRLYPSRCTSOHLNRTVTDNMLCAGDTRSGGPQANLHDA 300

Qy 323 CQGDGSGPLVCLNDGRMTLVGIIISWGLGCGQKDVPGVYTKVTNYLDWIRDNRMP 376
Db 301 CQGDGSGPLVCLNDGRMTLVGIIISWGLGCGQKDVPGVYTKVTNYLDWIRDNRMP 354

RESULT 10
US-09-974-298-145
; Sequence 145, Application US/09974298
; Patent No. US20020156263A1
; GENERAL INFORMATION:
; APPLICANT: Chen, Hwei-Mei
; TITLE OF INVENTION: GENES EXPRESSED IN BREAST CANCER
; FILE REFERENCE: PA-0037 P
; CURRENT APPLICATION NUMBER: US/09/974,298
; CURRENT FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: 60/238,331
; PRIOR FILING DATE: 2000-05-10
; NUMBER OF SEQ ID NOS: 194
; SOFTWARE: PERL Program
; SEQ ID NO 145
; LENGTH: 562
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20020156263A1 1001470CD1
US-09-974-298-145

Query Match          94.7%; Score 1952; DB 10; Length 562;
Best Local Similarity 99.4%; Pred. No. 6.5e-182;
Matches 354; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 21 AASEGNSDCYFGNGSAYRGTHSLTESGASCLPWNMILIGKVYTAQNPSAALGLGKHNY 80
```

```
Db 207 ACSEGNSDCYFGNGSAYRGTHSLTESGASCLRWNMILIGKVYTAQNPSAALGLGKHNY 266
Qy 81 CRNPDGDAKFWCHLVKNRRLTWECYDVPCSTCGLRQYQSQPQRIKGLFADIASHPWQAA 140
Db 267 CRNPDGDAKFWCHLVKNRRLTWECYDVPCSTCGLRQYQSQPQRIKGLFADIASHPWQAA 326
Qy 141 AIFAKHRRSPGERFLCGGILISSCWILSAAHCFQERPPPHLTVILGRYRVVPGEEQKEE 200
Db 327 AIFAKHRRSPGERFLCGGILISSCWILSAAHCFQERPPPHLTVILGRYRVVPGEEQKEE 386
Qy 201 FEYKYIVHKEFDDDDTYNDIALQLKSDSSRCAQESSVVRTVCLPPADLPDWTCELSG 260
Db 387 FEYKYIVHKEFDDDDTYNDIALQLKSDSSRCAQESSVVRTVCLPPADLPDWTCELSG 446
Qy 261 SGYGKHEALSPFYSERLKEAHVRLYPSRCTSOHLNRTVTDNMLCAGDTRSGGPQANLH 320
Db 447 SGYGKHEALSPFYSERLKEAHVRLYPSRCTSOHLNRTVTDNMLCAGDTRSGGPQANLH 506
Qy 321 DACQGDGSGPLVCLNDGRMTLVGIIISWGLGCGQKDVPGVYTKVTNYLDWIRDNRMP 376
Db 507 DACQGDGSGPLVCLNDGRMTLVGIIISWGLGCGQKDVPGVYTKVTNYLDWIRDNRMP 562

RESULT 11
US-09-987-457-14
; Sequence 14, Application US/09987457
; Publication No. US20030013150A1
; GENERAL INFORMATION:
; APPLICANT: Manosroi, Aranya
; APPLICANT: Manosroi, Jiradej
; APPLICANT: Tayapiwatana, Chatchai
; APPLICANT: Goetz, Friedrich
; APPLICANT: Warner, Rolf-Guenther
; TITLE OF INVENTION: Methods for Large Scale Protein Production in Prokaryotes
; FILE REFERENCE: 0652.2180001
; CURRENT APPLICATION NUMBER: US/09/987,457
; CURRENT FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: 60/268,573
; PRIOR FILING DATE: 2001-02-15
; PRIOR APPLICATION NUMBER: GB 00 27 782.2
; PRIOR FILING DATE: 2000-11-14
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 14
; LENGTH: 343
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: part of the
; OTHER INFORMATION: recombinant K2S molecule (modified)
US-09-987-457-14

Query Match          91.0%; Score 1875.5; DB 11; Length 343;
Best Local Similarity 96.9%; Pred. No. 9.7e-175;
Matches 343; Conservative 0; Mismatches 0; Indels 11; Gaps 1;

Qy 23 SEGNSDCYFGNGSAYRGTHSLTESGASCLPWNMILIGKVYTAQNPSAALGLGKHNYCR 82
Db 1 SEGNSD-----THSLTESGASCLPWNMILIGKVYTAQNPSAALGLGKHNYCR 49

Qy 83 NPDGDAPKFWCHLVKNRRLTWECYDVPCSTCGLRQYQSQPQRIKGLFADIASHPWQAAI 142
Db 50 NPDGDAPKFWCHLVKNRRLTWECYDVPCSTCGLRQYQSQPQRIKGLFADIASHPWQAAI 109

Qy 143 FAKHRRSPGERFLCGGILISSCWILSAAHCFQERPPPHLTVILGRYRVVPGEEQKEE 202
Db 110 FAKHRRSPGERFLCGGILISSCWILSAAHCFQERPPPHLTVILGRYRVVPGEEQKEE 169

Qy 203 VEKYIVHKEFDDDDTYNDIALQLKSDSSRCAQESSVVRTVCLPPADLPDWTCELSG 262
Db 170 VEKYIVHKEFDDDDTYNDIALQLKSDSSRCAQESSVVRTVCLPPADLPDWTCELSG 229
```



```
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 16
; LENGTH: 343
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: K2S 191-527,
; OTHER INFORMATION: modified
; US-09-987-455-16

Query Match          90.5%; Score 1865.5; DB 11; Length 343;
Best Local Similarity 96.6%; Pred. No. 9.2e-174;
Matches 342; Conservative 0; Mismatches 1; Indels 11; Gaps 1;

Qy 23 SEGSDCYFNGSAYRGTHSLTESGASCLPWNMILIGKVYTAQNPQAALGLGKHNCR 82
Db 1 SEGNSD-----THSLTESGASCLPWNMILIGKVYTAQNPQAALGLGKHNCR 49
Qy 83 NPDGDAKPWCHLVKNRRLTWECYCDVPSCTCGLRQYSPQPRIKGGLFADIASHPWQAAI 142
Db 50 NPDGDAKPWCHLVKNRRLTWECYCDVPSSTCGLRQYSPQPRIKGGLFADIASHPWQAAI 109
Qy 143 FAKHRRSPGERFLCGGILISSCWILSAAHCFQERFPPHLLTVILGRYRVVPGEEQKFE 202
Db 110 FAKHRRSPGERFLCGGILISSCWILSAAHCFQERFPPHLLTVILGRYRVVPGEEQKFE 169
Qy 203 VEKYIVHKEFDDTDYNDIALQLKSDSSRCAQESSVVRTVCLPPADLQLPDWTCELSG 262
Db 170 VEKYIVHKEFDDTDYNDIALQLKSDSSRCAQESSVVRTVCLPPADLQLPDWTCELSG 229
Qy 263 YGKHEALSPFYSERLKEAHVRLYPSRCTSQHLNRTVTDNMLCAGDTRSGGPOANLHDA 322
Db 230 YGKHEALSPFYSERLKEAHVRLYPSRCTSQHLNRTVTDNMLCAGDTRSGGPOANLHDA 289
Qy 323 CQDSSGGLVCLNDGRMTLVGIIISWGLCGGQKQDVPVYTKVTNYLDWIRDMRP 376
Db 290 CQDSSGGLVCLNDGRMTLVGIIISWGLCGGQKQDVPVYTKVTNYLDWIRDMRP 343

RESULT 15
US-09-987-457-12
; Sequence 12, Application US/09987457
; Publication No. US20030013150A1
; GENERAL INFORMATION:
; APPLICANT: Manosroi, Aranya
; APPLICANT: Manosroi, Jiradej
; APPLICANT: Tavapiwatana, Chatchai
; APPLICANT: Goetz, Friedrich
; APPLICANT: Werner, Rolf-Guenther
; TITLE OF INVENTION: Methods for Large Scale Protein Production in Prokaryotes
; FILE REFERENCE: 0652.2180001
; CURRENT APPLICATION NUMBER: US/09/987,457
; CURRENT FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: 60/268,573
; PRIOR FILING DATE: 2001-02-15
; PRIOR APPLICATION NUMBER: GB 00 27 782.2
; PRIOR FILING DATE: 2000-11-14
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 339
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: part of the
; OTHER INFORMATION: recombinant K2S molecule (modified)
; US-09-987-457-12

Query Match          89.8%; Score 1850.5; DB 11; Length 339;
Best Local Similarity 95.8%; Pred. No. 2.6e-172;
Matches 339; Conservative 0; Mismatches 0; Indels 15; Gaps 1;

Qy 23 SEGSDCYFNGSAYRGTHSLTESGASCLPWNMILIGKVYTAQNPQAALGLGKHNCR 82
```

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 8, 2003, 17:12:22 ; Search time 19 Seconds  
(without alignments)  
1908.189 Million cell updates/sec

Title: US-09-987-455-8  
Perfect score: 261  
Sequence: 1 MKKTAIAVALAGFATVAQ.....GVYTKVTNYLDWRDNRPG 377

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues  
Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 76:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1961	95.1	562	1 UKHUT	t-plasminogen acti
2	1638.5	79.5	559	1 A35029	t-plasminogen acti
3	1600	77.6	559	1 A29941	t-plasminogen acti
4	1451.5	70.4	394	2 JS0600	t-plasminogen acti
5	1428	69.3	477	2 JS0598	t-plasminogen acti
6	1427	69.2	431	2 JS0599	t-plasminogen acti
7	1423	69.0	477	1 A34369	t-plasminogen acti
8	1389.5	67.4	477	2 JS0597	t-plasminogen acti
9	779.5	37.8	433	1 UKMS	u-plasminogen acti
10	777	37.7	431	1 UKHU	u-plasminogen acti
11	774.5	37.6	433	1 UKBAY	u-plasminogen acti
12	772.5	37.5	432	1 S18932	u-plasminogen acti
13	769.5	37.3	442	1 UKPG	u-plasminogen acti
14	762	37.0	434	1 A35005	u-plasminogen acti
15	745	36.1	433	1 JN0560	u-plasminogen acti
16	738	35.8	655	1 A46688	hepatocyte growth
17	709.5	34.4	560	1 JC4795	plasma hyaluronan-
18	698.5	33.9	558	2 JC5878	plasma hyaluronan-
19	671.5	32.6	615	1 KFHU12	coagulation factor
20	670	32.5	603	2 S28941	coagulation factor
21	620.5	30.1	593	2 S45281	coagulation factor
22	600	29.1	812	1 PLMS	plasmin EC 3.4.21
23	580.5	28.2	460	2 B61545	plasmin EC 3.4.21
24	563.5	27.3	810	2 B30848	plasmin EC 3.4.21
25	561.5	27.2	810	1 PLHU	plasmin EC 3.4.21
26	552.5	26.8	790	1 PLPG	plasmin EC 3.4.21
27	550.5	26.7	812	1 PLBO	plasmin EC 3.4.21
28	547	26.5	810	2 I46260	plasmin EC 3.4.21
29	541.5	26.3	4548	1 S00657	apoprotein(a) (EC

RESULT 1

UKHUT

t-plasminogen activator (EC 3.4.21.68) precursor [validated] - human  
N/Alternate names: t-PA; tissue plasminogen activator

C/Species: Homo sapiens (man)

C/Date: 14-Nov-1983 #sequence revision 14-Nov-1983 #text change 08-Dec-2000

C/Accession: A94004; A23529; J00562; A93293; S02125; A91343; A93951; A91322; A54645; I60

R/Ny. T.; Elgh, F.; Lund, B.

Proc. Natl. Acad. Sci. U.S.A. 81, 5355-5359, 1984

A/Title: The structure of the human tissue-type plasminogen activator gene: correlation

A/Reference number: A94004; MUID:84298137; PMID:6089198

A/Accession: A94004

A/Molecule type: DNA

A/Residues: 1-562 <NVT>

A/Cross-references: GB:I00141

A/Note: The codon given for residue 93 (ACC) is inconsistent with the authors' translation.

R/Friezner Degen, S.J.; Rajput, B.; Reich, E.

J. Biol. Chem. 261, 6972-6985, 1986

A/Title: The human tissue plasminogen activator gene.

A/Reference number: A23529; MUID:86196143; PMID:3009482

A/Accession: A23529

A/Molecule type: DNA

A/Residues: 1-562 <DEG>

A/Cross-references: GB:K03021; NID:G339817; PIDN:AAA98809.1; PID:G339818

R/Itagaki, Y.; Yasuda, H.; Morinaga, T.; Mitsuoka, S.; Higashio, K.

Agric. Biol. Chem. 55, 1225-1232, 1991

A/Title: Purification and characterization of tissue plasminogen activator secreted by h

A/Reference number: J00562; MUID:91291340; PMID:1368681

A/Accession: J00562

A/Molecule type: mRNA

A/Residues: 31-562 <ITA>

A/Cross-references: DBJ:D01096; NID:G220128; PIDN:BA00881.1; PID:G441174

A/Experimental source: embryonic lung fibroblast IMR-90 cells

A/Note: part of this sequence, including the amino end of the mature protein, was confir

R/Pennica, D.; Holmes, W.E.; Kohr, W.J.; Harkins, R.N.; Vehar, G.A.; Ward, C.A.; Bennett

Nature 301, 214-221, 1983

A/Title: Cloning and expression of human tissue-type plasminogen activator cDNA in Esche

A/Reference number: A93293; MUID:83115262; PMID:6337343

A/Accession: A93293

A/Molecule type: mRNA

A/Residues: 1-562 <PEN>

A/Cross-references: GB:I00141

A/Experimental source: melanoma cells

R/Sasaki, H.; Saito, Y.; Hayashi, M.; Otsuka, K.; Niwa, M.

Nucleic Acids Res. 16, 5695, 1988

A/Title: Nucleotide sequence of the tissue-type plasminogen activator cDNA from human fe

A/Reference number: S02125; MUID:88262579; PMID:3133640

A/Accession: S02125

A/Status: translation not shown

A/Molecule type: mRNA

A/Residues: 1-562 <SAS>

A/Cross-references: EMBL:X07393; NID:G37243; PIDN:CAA30302.1; PID:G37244

C;Comment: Cleavage by plasmin or trypsin produces two chains held together by a single C;Comment: t-PA converts plasminogen to plasmin by hydrolyzing a single Arg-Val bond. If C;Comment: t-PA binds chain A of fibrin by kringle 2 and the fibronectin type I repeat. If C;Genetics:

A;Gene: GDB:PLAT

A;Cross-references: GDB:119496; OMIM:173370

A;Map position: 3p12-3p12

A;Introns: 24/3; 85/1; 122/1; 180/2; 211/1; 268/2; 297/1; 362/2; 408/1; 455/3; 510/3

C;Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat homology; fibrinolysis; glycoprotein; hydrolase; kringle; plasma; serine proteinase

C;Keywords: fibrinolysis; signal sequence #status predicted <SIG>

F;1-23/Domain: signal sequence #status predicted <SIG>

F;24-32/Domain: propeptide #status predicted <PRO>

F;33-562/Product: t-plasminogen activator #status experimental <MAT>

F;33-562/Product: t-plasminogen activator chain A #status experimental <ACH>

F;41-78/Domain: fibronectin type I repeat homology <IF1>

F;86-119/Domain: EGF homology <EGF>

F;127-208/Domain: kringle homology <KR1>

F;215-296/Domain: kringle homology <KR2>

F;311-562/Product: t-plasminogen activator chain B #status experimental <BCH>

F;311-562/Product: t-plasminogen activator chain B #status experimental <BCH>

F;41-71,69-78,86-97,91-108,110-119,127-208,148-190,179-203,215-296,236-278,267-291,299-4

F;152,483/Binding site: carbohydrate (Asn) (covalent) #status experimental

F;219/Binding site: carbohydrate (Asn) (covalent) (partial) #status experimental

F;310-311/Cleavage site: Arg-Ile (plasmin, trypsin) #status experimental

F;357,406/Cleavage site: His, Asp #status predicted

F;513/Active site: Ser #status experimental

Query Match 95.1%; Score 1961; DB 1; Length 562;

Best Local Similarity 99.7%; Pred. No. 3.1e-162;

Matches 355; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 21 AASEGNSDCYFGNGSAVYRGTHSLTESGASCLPWNMILIGKVYTAQNPSAQLGLGKHNY 80

Db 207 ACSEGNSDCYFGNGSAVYRGTHSLTESGASCLPWNMILIGKVYTAQNPSAQLGLGKHNY 266

Qy 81 CRNPDGDAKWPCHVLKNRRLTWECYDVPSCSTCGLRQVYSPQFRIKGGLFADTASHPWQA 140

Db 267 CRNPDGDAKWPCHVLKNRRLTWECYDVPSCSTCGLRQVYSPQFRIKGGLFADTASHPWQA 326

Qy 141 AIFAKHRRSPGERFLCGGILSSCWILSAACHCFQRPFPHHLTVILGRYRVVVPGEEOK 200

Db 327 AIFAKHRRSPGERFLCGGILSSCWILSAACHCFQRPFPHHLTVILGRYRVVVPGEEOK 386

Qy 201 FEVEKYIVHKEFDDTDYNDIALQLKSDSRCAQESSVVRVTCVLPADLQLPDWTECEL 260

Db 387 FEVEKYIVHKEFDDTDYNDIALQLKSDSRCAQESSVVRVTCVLPADLQLPDWTECEL 446

Qy 261 SGYKHEALSPFYSERLKEAHVRLYPSSRCTSQHLLNRTVTNMLCAGDTRSGGPOANLH 320

Db 447 SGYKHEALSPFYSERLKEAHVRLYPSSRCTSQHLLNRTVTNMLCAGDTRSGGPOANLH 506

Qy 321 DACOGDSGGLPLVCLNDGRTMLVGIISWGLGCGQKDPGVYTKVTNYLDIWRNMRP 376

Db 507 DACOGDSGGLPLVCLNDGRTMLVGIISWGLGCGQKDPGVYTKVTNYLDIWRNMRP 562

RESULT 2

A35029

t-plasminogen activator (EC 3.4.21.68) precursor - rat

C;Species: Rattus norvegicus (Norway rat)

C;Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999

C;Accession: A35029; A31597

R;Peng, P.; Ohlsson, M.; Ny, T.

J. Biol. Chem. 285, 2022-2027, 1990

A;Title: The structure of the TATA-less rat tissue-type plasminogen activator gene. Spec

A;Reference number: A35029; MUID:90130448; PMID:2105315

A;Accession: A35029

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-559 <FEN>

A;Cross-references: GB:M31197; NID:g207429; PIDN:AAA42261.1; PID:g207431; GB:J05226

R;Ny, T.; Leonardsson, G.; Hsueh, A.J.W.

DNA 7, 671-677, 1988

A>Title: Cloning and characterization of a cDNA for rat tissue-type plasminogen activator  
 A;Reference number: A31597; MUID:89170114; PMID:3148445  
 A;Accession: A31597  
 A;Molecule type: mRNA  
 A;Residues: 1-379; 'K', 381-559 <NVT>  
 A;Cross-references: GB:M23697; NID:g530159; PIDN:AAA41812.1; PID:g530160  
 C;Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat hom  
 C;Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase  
 F;1-17/Domain: signal sequence #status predicted <SIG>  
 F;18-29/Domain: propeptide #status predicted <PRO>  
 F;30-559/Product: t-plasminogen activator #status predicted <MAT>  
 F;30-308/Product: t-plasminogen activator chain A #status predicted <ACH>  
 F;38-75/Domain: fibronectin type I repeat homology <1F1>  
 F;83-116/Domain: EGF homology <EGF>  
 F;124-205/Domain: kringle homology <KR1>  
 F;213-294/Domain: kringle homology <KR2>  
 F;309-559/Product: t-plasminogen activator chain B #status predicted <BCH>  
 F;309-553/Domain: trypsin homology <TRY>  
 F;38-68,66-75,83-94,88-105,107-116,124-205,145-187,176-200,213-294,234-276,265-289,297-4  
 F;149,481/Binding site: carbohydrate (Asn) (covalent) #status predicted  
 F;308-309/Cleavage site: Arg-Ile (plasmin, trypsin) #status predicted  
 F;355,404,510/Active site: His, Asp, Ser #status predicted

Query Match 79.5%; Score 1638.5; DB 1; Length 559;  
 Best Local Similarity 82.5%; Pred. No. 3.1e-134; Mismatches 29; Indels 1; Gaps 1;  
 Matches 287; Conservative 29; Mismatches 31; Indels 1; Gaps 1;

QY 28 DCYFGNGSAYRGTHSLTESGASCLPWNMILIGKVYTAQNPSAQAALGLGKHNCRNPDGD 87  
 DB 212 DCYGVGVYRGTHSFTTSKASCLPWNMILIGKVYTAQNPSAQAALGLGKHNCRNPDGD 271

QY 88 AKPWCHVLKRRRLTWECYDVPCSTCGLRQYQSQPFRIKGLGFADIASHPQAAIFAKHR 147  
 DB 272 AKPWCHVMKDKRLTWECYDVPCSTCGLRQYQSQPFRIKGLGFADIASHPQAAIFVKNK 331

QY 148 RSPGRFLCGGLTSSCWLSAHCFOERFPPHLLTVILGRYRVVPGEEQKEVEKYI 207  
 DB 332 RSPGRFLCGGLTSSCWLSAHCFOERFPPHLLTVILGRYRVVPGEEQKEVEKYI 391

QY 208 VHKPEDDTDYNDIALQLKSDSSRCAQESSVVRVTCVLPADLPDWTCELSYGVKHE 267  
 DB 392 VHKPEDDTDYNDIALQLKSDSSRCAQESSVVRVTCVLPADLPDWTCELSYGVKHE 451

QY 268 ALSPPYERLKEAHLVLPSSRCTSQHLNRTVTDNMLCAGDTRSGGQPMNLHACQGD 327  
 DB 452 ASSPFSRLKEAHLVLPSSRCTSQHLNRTVTDNMLCAGDTRSGGQPMNLHACQGD 510

QY 328 GGPLVCLNDGRMTLVGIISWGLGCGQKQVPGYTKVTNYLDWIRDNR 375  
 DB 511 GGPLVCLNDGRMTLVGIISWGLGCGQKQVPGYTKVTNYLDWIRDNR 558

## RESULT 3

A29941  
 t-plasminogen activator (EC 3.4.21.68) precursor - mouse  
 C;Species: Mus musculus (house mouse)  
 C;Date: 10-Sep-1999 #sequence revision 10-Sep-1999 #text\_change 10-Sep-1999  
 C;Accession: A29941; S48205; S48207; S48206  
 R;Rickle, R.J.; Darrow, A.L.; Strickland, S.  
 J. Biol. Chem. 263, 1563-1569, 1988

A>Title: Molecular cloning of complementary DNA to mouse tissue plasminogen activator mR  
 A;Reference number: A29941; MUID:88087303; PMID:2826484  
 A;Accession: A29941  
 A;Molecule type: mRNA  
 A;Residues: 1-559 <RIC>  
 A;Cross-references: GB:J03520; NID:g202109; PIDN:AAA40470.1; PID:g202110  
 R;Liljgren, H.R.; van Hoef, B.; Beelen, V.; Collen, D.  
 Eur. J. Biochem. 224, 863-871, 1994

A>Title: Characterization of the murine plasma fibrinolytic system.  
 A;Reference number: S48202; MUID:95010076; PMID:7523120  
 A;Accession: S48205  
 A;Molecule type: protein  
 A;Residues: 33-37, 'X', 39-40 <LIU>

A;Accession: S48207  
 A;Molecule type: protein  
 A;Residues: 309-316 <LI2>  
 A;Accession: S48206  
 A;Molecule type: protein  
 A;Residues: 33-37, 'X', 39-40 <LIU>  
 C;Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat hom  
 C;Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase  
 F;1-17/Domain: signal sequence #status predicted <SIG>  
 F;18-29/Domain: propeptide #status predicted <PRO>  
 F;30-559/Product: t-plasminogen activator #status predicted <MAT>  
 F;30-308/Product: t-plasminogen activator chain A #status predicted <ACH>  
 F;38-75/Domain: fibronectin type I repeat homology <1F1>  
 F;83-116/Domain: EGF homology <EGF>  
 F;124-205/Domain: kringle homology <KR1>  
 F;213-294/Domain: kringle homology <KR2>  
 F;309-559/Product: t-plasminogen activator chain B #status predicted <BCH>  
 F;309-553/Domain: trypsin homology <TRY>  
 F;38-68,66-75,83-94,88-105,107-116,124-205,145-187,176-200,213-294,234-276,265-289,297-4  
 F;149,481/Binding site: carbohydrate (Asn) (covalent) #status predicted  
 F;308-309/Cleavage site: Arg-Ile (plasmin, trypsin) #status predicted  
 F;355,404,510/Active site: His, Asp, Ser #status predicted

Query Match 77.6%; Score 1600; DB 1; Length 559;

Best Local Similarity 79.5%; Pred. No. 6.8e-131;

Matches 283; Conservative 34; Mismatches 37; Indels 2; Gaps 2;

QY 21 AASGNS-DCVFGNGSAYRGTHSLTESGASCLPWNMILIGKVYTAQNPSAQAALGLGKHN 79

DB 204 ACPKGSDECVGKGYRGTHSLTTSQASCLPWNMILIGKVYTAQNPSAQAALGLGKHN 263

QY 80 YCRNPDGAPWCHVLKRRRLTWECYDVPCSTCGLRQYQSQPFRIKGLGFADIASHPWQ 139

DB 264 YCRNPDGAPWCHVLKRRRLTWECYDVPCSTCGLRQYQSQPFRIKGLGFADIASHPWQ 323

QY 140 AAIKAKRRSPGRFLCGGLTSSCWLSAHCFOERFPPHLLTVILGRYRVVPGEEQ 199

DB 324 AAIKAKRRSPGRFLCGGLTSSCWLSAHCFOERFPPHLLTVILGRYRVVPGEEQ 383

QY 200 KFEYKIVHKEPDDTDYNDIALQLKSDSSRCAQESSVVRVTCVLPADLPDWTCE 259

DB 384 TFEIKIVHEEFDDTDYNDIALQLKSDSSRCAQESSVVRVTCVLPADLPDWTCE 443

QY 260 LSGYKHEALSPFYERLKEAHLVLPSSRCTSQHLNRTVTDNMLCAGDTRSGGQPMNL 319

DB 444 LSGYKHEALSPFYERLKEAHLVLPSSRCTSQHLNRTVTDNMLCAGDTRSGGQPMNL 502

QY 320 HDACQSGGGLVCLNDGRMTLVGIISWGLGCGQKQVPGYTKVTNYLDWIRDNR 375

DB 503 HDACQSGGGLVCLNDGRMTLVGIISWGLGCGQKQVPGYTKVTNYLDWIRDNR 558

## RESULT 4

JS0600  
 t-plasminogen activator (EC 3.4.21.68) gamma precursor - common vampire bat  
 N;Alternate names: tissue plasminogen activator  
 C;Species: Desmodus rotundus (common vampire bat)  
 C;Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 16-Jul-1999  
 C;Accession: JS0600  
 R;Kraetzschmar, J.; Haendler, B.; Langer, G.; Boisdol, W.; Bringmann, P.; Alagon, A.; Don  
 Gene 105, 229-237, 1991

A>Title: The plasminogen activator family from the salivary gland of the vampire bat Des  
 A;Reference number: JS0597; MUID:92039036; PMID:1937019  
 A;Accession: JS0600  
 A;Molecule type: mRNA  
 A;Residues: 1-394 <KRA>  
 A;Cross-references: GB:M63990; NID:g166078; PIDN:AAA31595.1; PID:g166079  
 A;Note: The authors translated the codon ATC for residue 75 as Thr  
 C;Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat hom  
 C;Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase  
 F;1-21/Domain: signal sequence #status predicted <SIG>  
 F;22-36/Domain: propeptide #status predicted <PRO>  
 F;37-394/Product: plasminogen activator gamma #status predicted <PLA>

F:45-126/Domain: kringle homology <KRG>  
F:143-388/Domain: trypsin homology <TRY>  
F:45-126,66-108,97-121,131-262,174-190,182-251,276-351,308-324,341-369/Disulfide bonds:  
F:142-143/Cleavage site: His-Ser (plasma) #status predicted  
F:189,238,345/Active site: His, Asp, Ser #status predicted  
F:315/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 70.4%; Score 1451.5; DB 2; Length 394;  
Best Local Similarity 67.6%; Pred. No. 3.5e-118;  
Matches 263; Conservative 42; Mismatches 69; Indels 15; Gaps 2;

QY 3 KTAIAIALAG-----FATVQAASEGSDCYFGNGSAYRGTHSLTESGAS 49  
DB 6 KTKLLCVLLCGAVPSLPRQETRYQLARGSRAYGDPHATCYKDQGVYRGTWSTESGAQ 65  
QY 50 CLPWNMILIGKVTYQAQNSAALGLGKHNYCRNPDGAKPWCHVLKRRRLTWECVPS 109  
DB 66 CINWNSLLIRTYNGRMPAEVKLGLGHNHNYCRNPDGASKPCVYIKARKFTSBSCSV 125  
QY 110 CS--TCGLROYSQPQFRKGLFADIASHPWQAATFAKHRSRSPGERFLCGGILISSCWIL 167  
DB 126 CSKATCGLRKYKEPQLHSTGGLFTDITSHPWQAALFAQNRSSGERFLCGGILISSCWIL 185  
QY 168 SAAHCFQERFPFHLLTVILGRYVVPGESEKFEVEKYIVHKEFDDDDTYNDIALQLK 227  
DB 186 TAAHCFQERYPPQHLRVVLGRYVVKPKESQTFEVEKCIHVEEFDDDDTYNDIALQLK 245  
QY 228 SSSRCAQESSVVRTVCLPPADLQLPDWTCELSGYGKHEALSPYSERLKEAHVRLYPS 287  
DB 246 SGSPQCAQESDSVRAICLPEANLQLPDWTCELSGYGKHSSSPYSERLKEAHVRLYPS 305  
QY 288 SRTSOHLNRTVTNNMLCAGDTRSGGQANLHDCQSGPLVCLNDGRTVLGIIISW 347  
DB 306 SRTSKFLFNKVTNNMLCAGDTRSGEYVNVHDCQSGPLVCLNDGRTVLGIIISW 365  
QY 348 GLGCGQKQVPGVYTKVNTYLDWIRNMRP 376  
DB 366 GVGCGEKDIPGYTKVNTYLVGWRNMRP 394

RESULT 5  
JS0598  
t-plasminogen activator (EC 3.4.21.68) alpha-2 precursor - common vampire bat  
N:Alternate names: tissue plasminogen activator  
C:Species: Desmodus rotundus (common vampire bat)  
C>Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 16-Jul-1999  
C:Accession: JS0598  
R:Kraetzschmar, J.; Haendler, B.; Langer, G.; Boidol, W.; Bringmann, P.; Alagon, A.; Don  
Gene 105, 229-237, 1991  
A:Title: The plasminogen activator family from the salivary gland of the vampire bat Des  
A:Reference number: JS0597; MUID:92039036; PMID:1937019  
A:Accession: JS0598  
A:Molecule type: mRNA  
A:Residues: 1-477 <KRA>  
A:Cross-references: GB:M63988; NID:gl66074; PID:AAA31593.1; PID:gl66075  
C:Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat hom  
C:Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase  
F:1-21/Domain: signal sequence #status predicted <SIG>  
F:22-36/Domain: propeptide #status predicted <PRO>  
F:37-477/Product: plasminogen activator alpha-2 #status predicted <PLA>  
F:42-79/Domain: fibronectin type I repeat homology <IFA>  
F:87-120/Domain: EGF homology <EGF>  
F:128-209/Domain: kringle homology <KRG>  
F:226-471/Domain: trypsin homology <TRY>  
F:42-72,70-98,92-109,111-120,128-209,149-191,180-204,214-345,257-273,265-334,359-4  
F:185,398/Binding site: carbohydrate (Asn) (covalent) #status predicted  
F:225-226/Cleavage site: His-Ser (plasma) #status predicted  
F:272,321,428/Active site: His, Asp, Ser #status predicted

Query Match 69.3%; Score 1428; DB 2; Length 477;  
Best Local Similarity 70.2%; Pred. No. 4.8e-116;  
Matches 254; Conservative 40; Mismatches 66; Indels 2; Gaps 1;

QY 17 TVAQAASEGSDCYFGNGSAYRGTHSLTESGASCLPWNMILIGKVTYQAQNSAALGLG 76  
DB 116 TGKQCEVDTHATCYKDQGVYRGTWSTESGAQINWNSLLTRTYNGRSDAITLGLG 175  
QY 77 KHNYCRNPDGAKPWCHVLKRRRLTWECVPS--TCGLROYSQPQFRKGLGFADIA 134  
DB 176 NHNYCRNPDNNSKPCVYIKARKFTSBSCSV--TCGLROYSQPQFRKGLGFADIA 235  
QY 135 SHPWQAALFAKHRSRSPGERFLCGGILISSCWILSAHCFQERFPFHLLTVILGRYVVP 194  
DB 236 SHPWQAALFAQNRSSGERFLCGGILISSCWILTAHCFQERYPPQHLRVVLGRYVVKP 295  
QY 195 GESEKFEVEKYIVHKEFDDDDTYNDIALQLKSDSRCAQESSVVRTVCLPPADLQLPD 254  
DB 296 GKSEQTEVEKCIHVEEFDDDDTYNDIALQLKSGSPQCAQESDSVRAICLPEANLQLPD 355  
QY 255 WTCELSGYGKHEALSPYSERLKEAHVRLYPSRRTSQHLLNRTVTNNMLCAGDTRSGG 314  
DB 356 WTCELSGYGKHSSSPYSERLKEAHVRLYPSRRTSKEFLFNKVTNNMLCAGDTRSGE 415  
QY 315 PQANLHDCQSGDGGPLVCLNDGRTVLGIIISWGLCGQKQVPGVYTKVNTYLDWIRNMR 374  
DB 416 IYPNVHDCQSGDGGPLVCLNDGRTVLGIIISWGLCGQKQVPGVYTKVNTYLDWIRNMR 475  
QY 375 RP 376  
DB 476 RP 477

RESULT 6  
JS0599  
t-plasminogen activator (EC 3.4.21.68) beta precursor - common vampire bat  
N:Alternate names: tissue plasminogen activator  
C:Species: Desmodus rotundus (common vampire bat)  
C>Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 16-Jul-1999  
C:Accession: JS0599  
R:Kraetzschmar, J.; Haendler, B.; Langer, G.; Boidol, W.; Bringmann, P.; Alagon, A.; Don  
Gene 105, 229-237, 1991  
A:Title: The plasminogen activator family from the salivary gland of the vampire bat Des  
A:Reference number: JS0597; MUID:92039036; PMID:1937019  
A:Accession: JS0599  
A:Molecule type: mRNA  
A:Residues: 1-431 <KRA>  
A:Cross-references: GB:M63989; NID:gl66076; PID:AAA31594.1; PID:gl66077  
C:Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat hom  
C:Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase  
F:1-21/Domain: signal sequence #status predicted <SIG>  
F:22-36/Domain: propeptide #status predicted <PRO>  
F:37-431/Product: plasminogen activator beta #status predicted <PLA>  
F:42-74/Domain: EGF homology <EGF>  
F:82-163/Domain: kringle homology <KRG>  
F:180-425/Domain: trypsin homology <TRY>  
F:41-52,46-63,65-74,82-163,103-145,134-158,168-299,211-227,219-288,313-388/Disulfide bon  
F:139,352/Binding site: carbohydrate (Asn) (covalent) #status predicted  
F:179-180/Cleavage site: His-Ser (plasma) #status predicted  
F:226,275,382/Active site: His, Asp, Ser #status predicted  
F:345-361,378-406/Disulfide bonds: #status predicted

Query Match 69.2%; Score 1427; DB 2; Length 431;  
Best Local Similarity 69.9%; Pred. No. 5.3e-116;  
Matches 253; Conservative 41; Mismatches 66; Indels 2; Gaps 1;

QY 17 TVAQAASEGSDCYFGNGSAYRGTHSLTESGASCLPWNMILIGKVTYQAQNSAALGLG 76  
DB 70 TGKQCEVDTHATCYKDQGVYRGTWSTESGAQINWNSLLTRTYNGRSDAITLGLG 129  
QY 77 KHNYCRNPDGAKPWCHVLKRRRLTWECVPS--TCGLROYSQPQFRKGLGFADIA 134  
DB 130 NHNYCRNPDNNSKPCVYIKARKFTSBSCSV--TCGLROYSQPQFRKGLGFADIA 189  
QY 135 SHPWQAALFAKHRSRSPGERFLCGGILISSCWILSAHCFQERFPFHLLTVILGRYVVP 194  
DB 190 SHPWQAALFAQNRSSGERFLCGGILISSCWILTAHCFQERYPPQHLRVVLGRYVVKP 249



```
QY 195 GEEBQKFEVKEKIVHKEFDYNDIALQLKSDSSCAQESSVVRVTCVLPADLQPD 254
DB 250 KKEEQTFEVEKIIHEBFDYNDIALQLKSGSPQCAQESDVRAICLPEANLQPD 309
QY 255 WTECELSGVGKHEALSPPYSERLKEAHLVLPSSRCTSOHLNRTVTNMLCAGDTRSGG 314
DB 310 WTECELSGVGKHKSSPPYSERLKEAHLVLPSSRCTSKFLFNKVTNMLCAGDTRSGE 369
QY 315 PQANLHDAQCGDSGGPLVCLNDGRMTLVGIISWGLGCGQKQDVPVYTKVTNYLDWIRDNM 374
DB 370 IYPNVHDAQCGDSGGPLVCLNDGRMTLVGIISWGLGCGQKQDVPVYTKVTNYLDWIRDNM 429
QY 375 RP 376
DB 430 RP 431

RESULT 7
A34369
t-plasminogen activator (EC 3.4.21.68) precursor - false vampire bat (Megaderma lyra)
C:Species: Megaderma lyra
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: A34369
R:Gardell, S.J.; Duong, L.T.; Diehl, R.E.; York, J.D.; Hare, T.R.; Register, R.B.; Jacob
J. Biol. Chem. 264, 17947-17952, 1989
A:Title: Isolation, characterization, and cDNA cloning of a vampire bat salivary plasmin
A:Reference number: A34369; MUID:90036867; PMID:2509450
A:Accession: A34369
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-477 <GAR>
A:Cross-references: GB:M63082; NID:g166080; PIDN:AAA31596.1; PID:g166081
C:Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat hom
C:Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase
F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-36/Domain: propeptide #status predicted <PRO>
F:37-477/Product: plasminogen activator #status predicted <PLA>
F:42-79/Domain: fibronectin type I repeat homology <1FA>
F:128-209/Domain: kringle homology <KRG>
F:226-471/Domain: trypsin homology <TRY>
F:42-72,70-79,87-98,92-109,111-120,128-209,149-191,180-204,214-345,257-273,265-334,359-4
F:272,321,428/Active site: His, Asp, Ser #status predicted

Query Match 69.0%; Score 1423; DB 1; Length 477;
Best Local Similarity 70.2%; Pred. No. 1.3e-115;
Matches 254; Conservative 38; Mismatches 69; Indels 2; Gaps 1;

QY 17 TVAQAASEGNSDCYFGNSAYRGTHSLTESGASCLPWNMSMLIGKVYTAQNPSAQLGLG 76
DB 116 TGKQCEVDTHATCYKQDQGVYRGYTWSTESGAQCINWNSLLTRTYNGRSDAITLGLG 175
QY 77 KHNCRNPDGAKPWCHLVKNRRLTWECYDVPCS--TCGLRQYVQPOFRIGKGLFADIA 134
DB 176 NHNCRNPDNSKWCYVYIKASKEFILEPCVPCVCSKATCGLRKYKEPQLHSTGGLFTDIT 235
QY 135 SHPWQAALFAKHRSRPGERFLCGGLILSSCWILSAACHFCQERFPHHLTVILGRTYRVP 194
DB 236 SHPWQAALFAQNRRSGRERFLCGGLILSSCWILTAACHFCQERYPPQHURVVLGRTYRVP 295
QY 195 GEEBQKFEVKEKIVHKEFDYNDIALQLKSDSSCAQESSVVRVTCVLPADLQPD 254
DB 296 KKEEQTFEVEKIVHEBFDYNDIALQLKSGSPQCAQESDVRAICLPEANLQPD 355
QY 255 WTECELSGVGKHEALSPPYSERLKEAHLVLPSSRCTSOHLNRTVTNMLCAGDTRSGG 314
DB 356 WTECELSGVGKHKSSPPYSERLKEAHLVLPSSRCTSKFLFNKVTNMLCAGDTRSGE 415
QY 315 PQANLHDAQCGDSGGPLVCLNDGRMTLVGIISWGLGCGQKQDVPVYTKVTNYLDWIRDNM 374
DB 416 IHPNVHDAQCGDSGGPLVCLNDGRMTLVGIISWGLGCGQKQDVPVYTKVTNYLDWIRDNM 475
```

```
QY 375 RP 376
DB 476 RP 477

RESULT 8
JS0597
t-plasminogen activator (EC 3.4.21.68) alpha-1 precursor - common vampire bat
N:Alternate names: tissue plasminogen activator
C:Species: Desmodus rotundus (common vampire bat)
C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 16-Jul-1999
C:Accession: JS0597
R:Kraetzschmar, J.; Haendler, B.; Langer, G.; Boidol, W.; Bringmann, P.; Alagon, A.; Don
Gene 105, 229-237, 1991
A:Title: The plasminogen activator family from the salivary gland of the vampire bat Des
A:Reference number: JS0597; MUID:92039036; PMID:1937019
A:Accession: JS0597
A:Molecule type: mRNA
A:Residues: 1-477 <KRA>
A:Cross-references: GB:M63987; NID:g166070; PIDN:AAA31591.1; PID:g166071
C:Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat hom
C:Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase
F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-36/Domain: propeptide #status predicted <PRO>
F:37-477/Product: plasminogen activator alpha-1 #status predicted <PLA>
F:42-79/Domain: fibronectin type I repeat homology <1FA>
F:128-209/Domain: kringle homology <KRG>
F:226-471/Domain: trypsin homology <TRY>
F:42-72,70-79,87-98,92-109,111-120,128-209,149-191,180-204,214-345,257-273,265-334,359-4
F:272,321,428/Active site: His, Asp, Ser #status predicted

Query Match 67.4%; Score 1389.5; DB 2; Length 477;
Best Local Similarity 69.2%; Pred. No. 1.1e-112;
Matches 252; Conservative 37; Mismatches 72; Indels 3; Gaps 2;

QY 13 AGFATVAQAASEGNSDCYFGNSAYRGTHSLTESGASCLPWNMSMLIGKVYTAQNPSAQA 72
DB 113 AGY-TGKCEVDTHATCYEGQVYRGYTWSTAESRVCINWNSLLTRTYNGRMPDAFN 171
QY 73 LGLGKHNCRNPDGAKPWCHLVKNRRLTWECYDVPCS--TCGLRQYVQPOFRIGKGLF 130
DB 172 LGLGHNCRNPNPAGPKPWCYVYKAGKFTSESCVPCVCSKATCGLRKYKEPQLHSTGGLF 231
QY 131 ADIASHPWQAALFAKHRSRPGERFLCGGLILSSCWILSAACHFCQERFPHHLTVILGRTY 190
DB 232 TDITSHPWQAALFAQNRRSGRERFLCGGLILSSCWILTAACHFCQESYLPDLQKVLGRTY 291
QY 191 RVVPGEERQKFEVKEKIVHKEFDYNDIALQLKSDSSCAQESSVVRVTCVLPADL 250
DB 292 RVVPGEERQKFEVKEKIVHKEFDYNDIALQLKSDSPQCAQESDVRAICLPEANL 351
QY 251 QLPDWTCELSGVGKHEALSPPYSERLKEAHLVLPSSRCTSOHLNRTVTNMLCAGDT 310
DB 352 QLPDWTCELSGVGKHKSSPPYSERLKEAHLVLPSSRCAPKFLFNKVTNMLCAGDT 411
QY 311 RSGGPQANLHDAQCGDSGGPLVCLNDGRMTLVGIISWGLGCGQKQDVPVYTKVTNYLDWI 370
DB 412 RSGEIYPNVHDAQCGDSGGPLVCLNDGRMTLVGIISWGLGCGQKQDVPVYTKVTNYLDWI 471
QY 371 RDNM 374
DB 472 RDNM 475

RESULT 9
UKMS
u-plasminogen activator (EC 3.4.21.73) precursor - mouse
C:Species: Mus musculus (house mouse)
C:Date: 30-Sep-1987 #sequence_revision 30-Sep-1987 #text_change 18-Jun-1999
C:Accession: A29420; A24615
```

R.Degen, S.J.F.; Heckel, J.L.; Reich, E.; Degen, J.L.  
 Biochemistry 26, 8270-8279, 1987  
 A>Title: The murine urokinase-type plasminogen activator gene.  
 A:Reference number: A29420; MUID:86163489; PMID:2831940  
 A:Accession: A29420  
 A:Molecule type: DNA  
 A:Residues: 1-433 <DE>  
 A:Cross-references: GB:M17922; NID:g202296; PIDN:AAA40539.1; PID:g202297  
 R:Belin, D.; Vassalli, J.D.; Combepine, C.; Godeau, F.; Nagamine, Y.; Reich, E.; Koher, Eur. J. Biochem. 148, 225-232, 1985  
 A>Title: Cloning, nucleotide sequencing and expression of cDNAs encoding mouse urokinase  
 A:Reference number: A24615; MUID:85179474; PMID:2985383  
 A:Accession: A24615  
 A:Molecule type: mRNA  
 A:Residues: 1-433 <BEL>  
 A:Cross-references: GB:X02389; NID:g55127; PIDN:CAA26231.1; PID:g55128  
 C:Genetics:  
 A:Introns: 19/3; 30/1; 66/1; 124/2; 155/1; 229/2; 279/3; 326/1; 375/3  
 C:Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; try  
 C:Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine proteinase  
 F:1-20/Domain: signal sequence #status predicted <SIG>  
 F:21-178/Product: urokinase-type plasminogen activator chain A #status predicted <ACH>  
 F:32-63/Domain: EGF homology <EGF>  
 F:71-152/Domain: kringle homology <KRG>  
 F:180-433/Product: urokinase-type plasminogen activator chain B #status predicted <BGH>  
 F:180-421/Domain: tryptsin homology <TRY>  
 F:169-301, 211-227, 219-290, 315-384, 347-363, 374-402/Disulfide bonds: #status predicted  
 F:226, 277, 378/Active site: His, Asp, Ser #status predicted

Query Match 37.8%; Score 779.5; DB 1; Length 433;  
 Best Local Similarity 40.3%; Pred. No. 8.6e-60;  
 Matches 148; Conservative 67; Mismatches 129; Indels 23; Gaps 5;

Qy	24	EGNSDCYCGNSAYRGTHSLTSGASCLPWSMILIGKVTYQNPQAQGLGKHNYCRN	83
Db	66	DASKTCYHGSDSVRGKANTDTKGRPCLAWNAVALQKPYNAHRPDAISLGLGKHNYCRN	125
Qy	84	PDGAKPWCHLVKNRRLTWECYDVPSCS-----TCGLROYSPQPRKGG	129
Db	126	PDNQKRPWCYQVIGLRFQVQEGWHDCLSKKPSVVDDQGFQCG-QALAPRFKIVGGE	184
Qy	130	FADTASHPWQAIAFAKRRSPGFLCGHILISSCWILSAACHQERPPPHHLTVILGRT	189
Db	185	FTEVENQWFAAIIYKNGKGGSPGFKCGSLSPCWVASAACHFQLPKKENYVVYLGQS	244
Qy	190	YRVVPGEEQFEVEKYIVHKEFDDT--YNDTALLQKSDSRCAQESSVWTVCLPP	247
Db	245	KESSYNPGEMKFEVEQLILHYREDLSLAYHNDIALKIRTSQCCAQPSRSIQITCLPP	304
Qy	248	ADLQLPDWTCEBLSGCGKHEALSPPYSERLKEAHVRLVPSRSCTSOHLNRTVTDNMLCA	307
Db	305	RFTDAPFGSDCEITGFGKESSDVLYPKNLKMSVVLVSHVQCQMPHYGSEINYMMLCA	364
Qy	308	GDRSGGQANLHADCGSDGGLVCLNDGRMTLVGIISWGLGCGQKQDVPGVYTKVTNYL	367
Db	365	AD-----PEWKT-DSCKGSDGGLICNIEGRPTLSGIVSGRGCAEKKNKGVYTRVSHFL	418
Qy	368	DWIRDNM 374	
Db	419	DWIOSHI 425	

RESULT 10  
 UKHU  
 u-plasminogen activator (EC 3.4.21.73) precursor [validated] - human  
 N:Alternate names: cellular plasminogen activator; urokinase; urokinase-type plasminogen  
 N:Contains: urokinase-type plasminogen activator chain A; urokinase-type plasminogen act  
 in form  
 C:Species: Homo sapiens (man)  
 C>Date: 17-Dec-1982 #sequence revision 04-Dec-1986 #text change 15-Sep-2000  
 C:Accession: A00931; I52209; JTO102; A37561; I38102; S65783; A37562; A37563; A37564; A35  
 R:Riccio, A.; Grimaldi, G.; Verde, P.; Sebastio, G.; Boast, S.; Blasi, F.  
 Nucleic Acids Res. 13, 2759-2771, 1985

A>Title: The human urokinase-plasminogen activator gene and its promoter.  
 A:Reference number: A00931; MUID:85215647; PMID:2987867  
 A:Accession: A00931  
 A:Molecule type: DNA  
 A:Residues: 1-431 <RIC>  
 A:Cross-references: GB:X02419; NID:g37601; PIDN:CAA26268.1; PID:g1834524  
 A>Note: the authors translated the codon ATG for residue 214 as Ile  
 R:Nagamine, Y.; Pearson, D.; Grattan, M.  
 Biochem. Biophys. Res. Commun. 132, 563-569, 1985  
 A>Title: Exon-intron boundary sliding in the generation of two mRNAs coding for porcine  
 A:Reference number: I52209; MUID:86050639; PMID:3933505  
 A:Accession: I52209  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 145-161 <NAG1>  
 A:Cross-references: GB:K03027; NID:g340174; PIDN:AAA61257.1; PID:g340175  
 R:Nagai, M.; Hiramatsu, R.; Kaneda, T.; Hayasuke, N.; Arimura, H.; Nishida, M.; Suyama, Gene 36, 183-188, 1985  
 A>Title: Molecular cloning of cDNA coding for human preprourokinase.  
 A:Reference number: JTO102; MUID:86056954; PMID:2415429  
 A:Accession: JTO102  
 A:Molecule type: mRNA  
 A:Residues: 1-213, 'I', 215-431 <NAG2>  
 A:Cross-references: GB:K03226; NID:g340155; PIDN:AA097138.1; PID:g340158; GB:DO0244; NID:R:Verde, P.; Stoppelli, M.P.; Galeffi, P.; Di Nocera, P.; Blasi, F.  
 Proc. Natl. Acad. Sci. U.S.A. 81, 4727-4731, 1984  
 A>Title: Identification and primary sequence of an unspliced human urokinase poly(A)+ RN  
 A:Reference number: A37561; MUID:84272706; PMID:6569620  
 A:Accession: A37561  
 A:Molecule type: mRNA  
 A:Residues: 66-431 <VER>  
 A:Cross-references: GB:DO0244; NID:g220138  
 R:Jacobs, P.; Cravador, A.; Loriau, R.; Brockly, F.; Colau, B.; Chuchana, P.; van Elsen, DNA 4, 139-146, 1985  
 A>Title: Molecular cloning, sequencing, and expression in Escherichia coli of human prep  
 A:Reference number: I38102; MUID:85203359; PMID:3888571  
 A:Accession: I38102  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-150, 'W', 152-213, 'I', 215-385, 'C', 387-429, 'V', 431 <JAC>  
 A:Cross-references: EMBL:X02760; NID:g35297; PIDN:CAA26535.1; PID:g35298  
 R:Yoshimoto, M.; Ushiyama, Y.; Sakai, M.; Tamaki, S.; Hara, H.; Takahashi, K.; Sawasaki, Biochim. Biophys. Acta 1293, 83-89, 1996  
 A>Title: Characterization of single chain urokinase-type plasminogen activator with a no  
 A:Reference number: S65783; MUID:56186279; PMID:8652631  
 A:Accession: S65783  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 21-140, 'L', 142-213, 'I', 215-431 <YOS>  
 A:Cross-references: EMBL:D11143; NID:g1311467; PIDN:BAA01919.1; PID:g1199928  
 R:Gunzler, W.A.; Steffens, G.J.; Otting, F.; Kim, S.M.A.; Frankus, E.; Flohe, L.  
 Hoppe-Seyler's Z. Physiol. Chem. 363, 1155-1165, 1982  
 A>Title: The primary structure of high molecular mass urokinase from human urine.  
 A:Reference number: A37562; MUID:83055084; PMID:6754569  
 A:Accession: A37562  
 A:Molecule type: protein  
 A:Residues: 21-177 <GUN>  
 R:Schaller, J.; Nick, H.; Rickli, E.E.; Gillesen, D.; Lergier, W.; Studer, R.O.  
 Eur. J. Biochem. 125, 251-257, 1982  
 A>Title: Human low-molecular-weight urinary urokinase. Partial characterization and prel  
 A:Reference number: A37563; MUID:83003608; PMID:6749491  
 A:Accession: A37563  
 A:Molecule type: protein  
 A:Residues: 136-176; 179-193, 'T', 195, 'T', 197-224 <SCH>  
 R:Steffens, G.J.; Gunzler, W.A.; Otting, F.; Frankus, E.; Flohe, L.  
 Hoppe-Seyler's Z. Physiol. Chem. 363, 1043-1058, 1982  
 A>Title: The complete amino acid sequence of low molecular mass urokinase from human uri  
 A:Reference number: A37564; MUID:83055099; PMID:6754572  
 A:Accession: A37564  
 A:Molecule type: protein  
 A:Residues: 158-410 <STE>  
 R:Kentzer, E.J.; Buko, A.; Menon, G.; Sarin, V.K.  
 Biochem. Biophys. Res. Commun. 171, 401-406, 1990

A;Title: Carbohydrate composition and presence of a fucose-protein linkage in recombinant  
A;Reference number: A35689; MUID:90365737; PMID:2393398  
A;Accession: A35689  
A;Molecule type: protein  
A;Residues: 21-30,'X',34-38,'X',40-43 <EN>  
A;Note: identification of a fucose and attempt to determine its attachment site  
R;Rabbani S.A.; Desjardins, J.; Bell, A.W.; Banville, D.; Mazar, A.; Henkin, J.; Goltz  
Biochem. Biophys. Res. Commun. 173, 1058-1064, 1990  
A;Title: An amino-terminal fragment of urokinase isolated from a prostate cancer cell line  
A;Reference number: A36697; MUID:91097529; PMID:2125213  
A;Accession: A36697  
A;Molecule type: protein  
A;Residues: 21-34 <RAB>  
R;Li, X.; Bokman, A.M.; Llinas, M.; Smith, R.A.G.; Dobson, C.M.  
submitted to the Brookhaven Protein Data Bank, July 1993  
A;Reference number: A51255; PDB:1KDU  
A;Contents: annotation; conformation and disulfide bond assignments by (1)H-NMR, residue  
R;Li, X.; Smith, R.A.G.; Dobson, C.M.  
Biochemistry 31, 9562-9571, 1992  
A;Title: Sequential (1)H NMR assignments and secondary structure of the kringle domain of  
A;Reference number: A44375; MUID:93003110; PMID:1327118  
A;Contents: annotation; conformation and disulfide bond assignments by (1)H-NMR  
R;Hansen, A.P.; Petros, A.M.; Meadows, R.P.; Nettekheim, D.G.; Mazar, A.P.; Olejniczak,  
submitted to the Brookhaven Protein Data Bank, January 1994  
A;Reference number: A66822; PDB:1URK  
A;Contents: annotation; conformation and disulfide bond assignments by (1)H-NMR, residue  
R;Spraggon, G.S.; Phillips, C.; Nowak, U.K.; Ponting, C.P.; Saunders, D.; Dobson, C.M.;  
submitted to the Brookhaven Protein Data Bank, July 1995  
A;Reference number: A66058; PDB:1LWV  
A;Contents: annotation; X-ray crystallography, 2.5 angstroms, residues 168-175;179-426  
C;Comment: This enzyme is found in urine in a high molecular mass form, consisting of A  
C;Comment: Urokinase-type plasminogen activator proteolytically activates plasminogen, a  
C;Genetics:  
A;Gene: GDB:PLAU  
A;Cross-references: GDB:119497; OMIM:191840  
A;Map position: 10q24-10q24  
A;Introns: 19/3; 29/1; 123/2; 154/1; 227/2; 277/1; 324/1; 373/3  
C;Function:  
A;Description: proteolytically activates plasminogen  
A;Pathway: fibrinolysis  
C;Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; try  
C;Keywords: fibrinolysis; glycoprotein; heterodimer; hydrolase; kringle; serine proteina  
F;1-20/Domain: signal sequence #status predicted <SIG>  
F;21-431/Product: urokinase-type plasminogen activator, single chain form #status predic  
F;21-177/Product: urokinase-type plasminogen activator chain A #status experimental <MP  
F;31-62/Domain: EGF homology <EGF>  
F;70-151/Domain: kringle homology <KRG>  
F;156-177/Product: urokinase-type plasminogen activator chain A1 #status experimental <M  
F;179-431/Product: urokinase-type plasminogen activator chain B #status experimental <M  
F;179-419/Domain: trypsin homology <TRY>  
F;31-39,33-51,53-62,70-151,91-133,122-146,168-299,209-225,217-288,313-382,345-361,372-40  
F;36/Binding site: carbohydrate (Thr) (covalent) #status predicted  
F;178-179/Cleavage site: Lys-Ile (plasmin) #status experimental  
F;224,275,376/Active site: His, Asp, Ser #status experimental  
F;322/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match	37.7%;	Score	777;	DB	1;	Length	431;	
Best Local Similarity	43.3%;	Pred.	No.	1.4e-59;				
Matches	157;	Conservative	48;	Mismatches	134;	Indels	24; Gaps	6;

Qy	29	CYFGNGSAYRGTHSLTESGASCLPWNSSMILIGKYVTTAQNPSAALGLGKHENYCRNPDGDA	88
Db	70	CYEGNGHFYRGKASTDTMGPRCLPWNSTAVLQQTYHAHRSDALQLGLGKHENYCRNPDNR	129
Qy	89	KPCHVLNKRRLTWECYDVPSCS-----TCGLQVSOQPFRIKGLGFADIA	134
Db	130	RPWCYVQVGLKPLVQECVHWDCADGKKPSSPBPBLKFCQC-QKTLRPRFXIIGGEFTTIE	188
Qy	135	SHPWQAAIFAKHRESPCRFLCCGILLSSCWILLSAAHCFQERFPHHLTVILGRTYRVP	194
Db	189	NQWFAAIYRRH-RGGSVTVVCGSLMSPCWVISAHCFIDYPKKEDIYVILGRSLNSN	247
Qy	195	GEEQKFEVRYKTVHKFDDDT--YNDNIALLOLKSDSSRCAOESSVVRVTCVLPADLQ	252

RESULT 12  
S18932  
u-plasminogen activator (EC 3.4.21.73) precursor - rat  
N:Alternate names: plasminogen activator, urokinase-type; urinary plasminogen activator  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 18-Oct-1989 #sequence revision 10-Feb-1995 #text\_change 18-Jun-1999  
C:Accession: S24604; I60186; I53472; S18932  
R:Rabbani, S.A.  
submitted to the EMBL Data Library, April 1992  
A:Reference number: S24604  
A:Accession: S24604  
A:Molecule type: mRNA  
A:Residues: 1-15, 'H', 17-23, 'G', 25-331, 'N', 333-432 <RAB>  
A:Cross-references: EMBL:X65851; NID:G57456; PIDN:CAA46601.1; PID:G57457  
R:Henderson, B.R.; Tansey, W.P.; Phillips, S.M.; Ramshaw, I.A.; Kefford, R.F.  
Cancer Res. 52, 2489-2496, 1992  
A:Title: Transcriptional and posttranscriptional activation of urokinase plasminogen act  
A:Reference number: I60186; MUID:922333409; PMID:1568219  
A:Accession: I60186  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-432 <RES>  
A:Cross-references: EMBL:X63434; NID:G57465; PIDN:CAA45028.1; PID:G57466  
A:Experimental source: strain Fischer 344; tissue mammary  
R:Ragno, P.; Cassano, S.; Degen, J.; Kessler, C.; Blasi, F.; Rossi, G.  
FEBS Lett. 306, 193-198, 1992  
A:Title: The receptor for the plasminogen activator of urokinase type is up-regulated in  
A:Reference number: I53472; MUID:92339549; PMID:1321734  
A:Accession: I53472  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 31-62 <R2>  
A:Cross-references: EMBL:X66907; NID:G396200; PIDN:CAA47356.1; PID:G938279  
C:Genetics:  
A:Gene: uPA  
C:Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; try  
C:Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine proteinase  
F:1-19/Domain: signal sequence #status predicted <SIG>  
F:20-177/Product: urokinase-type plasminogen activator chain A #status predicted <ACH>  
F:31-62/Domain: EGF homology <EGF>  
F:70-151/Domain: kringle homology <KR>  
F:179-432/Product: urokinase-type plasminogen activator chain B #status predicted <BCH>  
F:179-430/Domain: trypsin homology <TRY>  
F:168-300,210-226,218-289,314-383,346-362,373-401/Disulfide bonds: #status predicted  
F:225,276,377/Active site: His, Asp, Ser #status predicted  
Query Match 37.5%; Score 772.5; DB 1; Length 432;  
Best Local Similarity 40.6%; Pred. No. 3.5e-59;  
Matches 147; Conservative 64; Mismatches 128; Indels 23; Gaps 5;  
Qy 29 CYFGNGSAYRGTHSLTESGASCLPWNMILIGKYTAQNPSAALGLGKHNCRNPDGA 88  
Db 70 CYHNGQSYRGKANTDTGRFCLAWNSPAVLQQTNAHRSALSLGLGKHNCRNPDNR 129  
Qy 89 KPWCHVLKNRLTWECYCDVPSCS-----TCGLRQYSQPOFRKGLFADIA 134  
Db 130 RPWCYQVGLQVQECMVQDCSLSKPSSVTDDQGGFCG-QKALRPRFKVGGFTWE 188  
Qy 135 SHPQAAIAFAKRRSPGRFLCGGILISSCWILSAAHCFQERFPPHILTVILGRYRVP 194  
Db 189 NQPFVAAIYLNKGGSPSPFKCGSLISPCWVASATHCFVNPQKEEVVVLGSKNSY 248  
Qy 195 GEEQKFEVEKYIVHKEFDDT--YNDIALQLKSSRCAQSSSVVTVCLPPADLQL 252  
Db 249 NPGEMKFEVEQILIHEDFSDTLAFPHNDIALKIRTSTGCAQPSRTIQTICLPPFGDA 308  
Qy 253 PDWTECELSGKHEALSPFYSERLKEAHVRLYPSRCTSOHLNRRVTNDMLCAGDTR 312  
Db 309 PFGSDCEITGFGQSATDYFPKOLKMSVVKIISHQCKQPHYYSYSEINYMCLCAAD--- 365

Qy 313 GGPQANLHDACQDGGSGGLVCLNDGRMTLVGIIWSGLGCGQKDVGPVYTKVTNYLDWIRD 372  
Db 366 --PEWKT-DSCSGDGGGLICNIDGRPTLSGIWSGSGCAEKPKPGVYTVSVFLNWIQS 422  
Qy 373 NM 374  
Db 423 HI 424  
RESULT 13  
UKPG  
u-plasminogen activator (EC 3.4.21.73) precursor - pig  
N:Alternate names: uPA  
C:Species: Sus scrofa domestica (domestic pig)  
C:Date: 04-Dec-1986 #sequence\_revision 17-Mar-1987 #text\_change 07-Aug-1998  
C:Accession: A00932  
R:Nagamine, Y.; Pearson, D.; Altus, M.S.; Reich, E.  
Nucleic Acids Res. 12, 9525-9541, 1984  
A:Title: cDNA and gene nucleotide sequence of porcine plasminogen activator.  
A:Reference number: A00932; MUID:85087954; PMID:6096832  
A:Accession: A00932  
A:Molecule type: DNA  
A:Residues: 1-240, 'H', 242-442 <NAG1>  
A:Experimental source: kidney cell line LLC-PK1  
R:Nagamine, Y.  
submitted to the Protein Sequence Database, December 1986  
A:Reference number: A37566  
A:Contents: annotation; correction to residue 241  
C:Genetics:  
A:Introns: 19/3; 31/1; 67/1; 125/2; 165/2; 238/2; 288/1; 335/1; 384/3  
C:Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; try  
C:Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine proteinase  
F:1-20/Domain: signal sequence #status predicted <SIG>  
F:21-188/Product: urokinase-type plasminogen activator chain A #status predicted <ACH>  
F:33-64/Domain: EGF homology <EGF>  
F:72-153/Domain: kringle homology <KR>  
F:190-442/Product: urokinase-type plasminogen activator chain B #status predicted <BCH>  
F:150-430/Domain: trypsin homology <TRY>  
F:152/Binding site: carboxydrate (Asn) (covalent) #status predicted  
F:179-310,220-236,228-239,324-393,356-372,383-411/Disulfide bonds: #status predicted  
F:235,286,387/Active site: His, Asp, Ser #status predicted  
Query Match 37.3%; Score 769.5; DB 1; Length 442;  
Best Local Similarity 41.2%; Pred. No. 6.5e-59;  
Matches 153; Conservative 50; Mismatches 135; Indels 33; Gaps 6;  
Qy 29 CYFGNGSAYRGTHSLTESGASCLPWNMILIGKYTAQNPSAALGLGKHNCRNPDGA 88  
Db 72 CFEGNGHSYRGKANTTGTGRPCLPWNSATVLLNTYHAHRPDALQLGLGKHNCRNPDNR 131  
Qy 89 KPWCHVLKNRLTWECYCDVPSCS-----TCGLRQYSQPOFRK 125  
Db 132 RPWCYQVGLQVQECMVQDCSWPNCSSGSHRPAYDGKMPFSTPEKVFQCG-QKALRPRFKI 190  
Qy 126 KGLFLADIAHPQAAIAFAKRRSPGRFLCGGILISSCWILSAAHCFQERFPPHILTVI 185  
Db 191 VGGKSTTIENQPFVAAIYRHH-RGGSVTVYCGGSLISPCWVASATHCFVNPQKEDIVY 249  
Qy 186 LGTRYRVVPGEEQKFEVEKYIVHKEFDDT--YNDIALQLKSSRCAQSSSVVTV 243  
Db 250 LGRQTLHSSTHGEMKFEVEKLIHEDYSADSLAHNDIALKIRTDKGCAQPSRSIQT 309  
Qy 244 CLPADLQLPDWTECELSGKHEALSPFYSERLKEAHVRLYPSRCTSOHLNRRVTND 303  
Db 310 CLPFPVNGDAHFGASCETVGGKEDPDDYLPQGLKMTVVKLVSHRECCQPHYYSVETTK 369  
Qy 304 MLCAGDTRSGGPOANLHDACQDGGSGGLVCLNDGRMTLVGIIWSGLGCGQKDVGPVYTKV 363  
Db 370 MLCAD-----PQWKT-DSQCGDGGGLVCLSTGRLTLTGIVSGRECAMKDEKGVYTRV 423  
Qy 364 TNYLDWIRDM 374  
Db 424 SRELTWITHV 434





GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 8, 2003, 17:10:32 ; Search time 15 Seconds  
(without alignments)  
1181.938 Million cell updates/sec

Title: US-09-987-455-8

Perfect score: 2061

Sequence: 1 MKKTAIAVALAGFATVAQ.....GVYTKVTNYLDWIRNMRPG 377

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_41.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match	Length	DB	ID	Description
1	1961	95.1	562	1	TPA_HUMAN	P00750	homo sapien
2	1660	80.5	566	1	TPA_BOVIN	Q28198	bos taurus
3	1638.5	79.5	559	1	TPA_RAT	P19637	rattus norv
4	1601	77.7	559	1	TPA_MOUSE	P11214	mus musculus
5	1451.5	70.4	394	1	URT2_DESRO	P49150	desmodus ro
6	1428	69.3	477	1	URT2_DESRO	P15638	desmodus ro
7	1427	69.2	431	1	URT2_DESRO	P38121	desmodus ro
8	1389.5	67.4	477	1	URT1_DESRO	P98119	desmodus ro
9	779.5	37.8	433	1	UROK_MOUSE	P06869	mus musculus
10	777	37.7	431	1	UROK_HUMAN	P00749	homo sapien
11	774.5	37.6	433	1	UROK_PAPCY	P16227	papio cynoc
12	772.5	37.5	432	1	UROK_RAT	P29598	rattus norv
13	769.5	37.3	442	1	UROK_PIG	P04185	sus scrofa
14	762	37.0	434	1	UROK_CHICK	P15120	gallus gall
15	745	36.1	433	1	UROK_BOVIN	Q05589	bos taurus
16	742.5	36.0	653	1	HGFA_MOUSE	Q9r098	mus musculus
17	738	35.8	655	1	HGFA_HUMAN	Q04756	homo sapien
18	671.5	32.6	615	1	FA12_HUMAN	P00748	homo sapien
19	670	32.5	603	1	FA12_CAVPO	Q04962	cavia porce
20	620.5	30.1	593	1	FA12_BOVIN	P29140	bos taurus
21	607	29.5	812	1	PLMN_MOUSE	P20918	mus musculus
22	573.5	27.8	333	1	PLMN_CANFA	P80009	canis fami
23	571.5	27.7	343	1	PLMN_SHEEP	P81286	ovis aries
24	563.5	27.3	810	1	PLMN_MACMU	P12545	macaca mula
25	561.5	27.2	810	1	PLMN_HUMAN	P00747	homo sapien
26	552.5	26.8	790	1	PLMN_PIG	P06867	sus scrofa
27	550.5	26.7	812	1	PLMN_BOVIN	P06868	bos taurus
28	547	26.5	810	1	PLMN_ERIEU	Q39485	erinaceus e
29	541.5	26.3	4548	1	APOA_HUMAN	P08519	homo sapien
30	538.5	26.1	338	1	PLMN_HORSE	P80010	equus cabal
31	519	25.2	1420	1	APOA_MACMU	P14417	macaca mula
32	469.5	22.8	875	1	NETR_HUMAN	P56730	homo sapien
33	465	22.6	761	1	NETR_MOUSE	O08762	mus musculus

34	449	21.8	855	1	ST14_HUMAN	Q9y5y6	homo sapien
35	438	21.3	855	1	ST14_MOUSE	P56677	mus musculus
36	408	19.8	455	1	TMS5_MOUSE	Q9er04	mus musculus
37	404	19.6	490	1	TMS2_MOUSE	Q9jiq8	mus musculus
38	403	19.6	457	1	TMS5_HUMAN	Q9h3a3	homo sapien
39	402	19.5	418	1	HATT_HUMAN	O60235	homo sapien
40	399	19.4	603	1	CFAT_MOUSE	Q51129	mus musculus
41	398.5	19.3	275	1	TRYT_PIG	Q9n2d1	sus scrofa
42	398.5	19.3	625	1	THRB_BOVIN	P00735	bos taurus
43	396.5	19.2	261	1	KLK1_MOUSE	P15947	mus musculus
44	396.5	19.2	618	1	THRB_MOUSE	P19221	mus musculus
45	394	19.1	604	1	CFAT_RAT	Q9wuw3	rattus norv

ALIGNMENTS

RESULT 1				
TPA_HUMAN	ID	TPA_HUMAN	STANDARD;	PRT; 562 AA.
AC	P00750; Q15103;			
DT	21-JUL-1986 (Rel. 01, Created)			
DT	21-JUL-1986 (Rel. 01, Last sequence update)			
DT	15-SEP-2003 (Rel. 42, Last annotation update)			
DE	Tissue-type plasminogen activator precursor (EC 3.4.21.68) (tPA)			
DE	(t-PA) (t-plasminogen activator) (Alteplase) (Reteplase).			
GN	PLAT.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Melanoma;			
RA	MEDLINE=83115262; PubMed=6337343;			
RA	Pennica D., Holmes W.E., Kohr W.J., Harkins R.N., Vehar G.A.,			
RA	Ward C.A., Bennett W.F., Yelverton E., Seeburg P.H., Heyneker H.L.,			
RA	Goeddel D.V., Collen D.;			
RT	"Cloning and expression of human tissue-type plasminogen activator			
RT	cDNA in E. coli.;"			
RL	Nature 301:214-221(1983).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Fetal lung;			
RX	MEDLINE=88262579; PubMed=3133640;			
RA	Sasaki H., Saito Y., Hayashi M., Otsuka K., Niwa M.;			
RT	"Nucleotide sequence of the tissue-type plasminogen activator cDNA			
RT	from human fetal lung cells.;"			
RL	Nucleic Acids Res. 16:5695-5695(1988).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=88054470; PubMed=2824147;			
RA	Reddy V.B., Garramone A.J., Sasak H., Wei C.-M., Watkins P., Galli			
RA	Hsiung N.;			
RT	"Expression of human uterine tissue-type plasminogen activator in			
RT	mouse cells using BPV vectors.;"			
RL	DNA 6:461-472(1987).			
RN	[4]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=86196143; PubMed=3009482;			
RA	Friezner Degen S.J., Rajput B., Reich E.;			
RT	"The human tissue plasminogen activator gene.;"			
RL	J. Biol. Chem. 261:6972-6985(1986).			
RN	[5]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=84298137; PubMed=6089198;			
RA	Ny T., Elgh F., Lund B.;			
RT	"The structure of the human tissue-type plasminogen activator gene			
RT	correlation of intron and exon structures to functional and			
RT	structural domains.;"			
RL	Proc. Natl. Acad. Sci. U.S.A. 81:5355-5359(1984).			
RN	[6]			
RP	SEQUENCE FROM N.A.			

RA MEDLINE=86284200; PubMed=3090401;  
RA Harris T.J., Patel T., Marston F.A., Little S., Entage J.S.,  
RA Opdenakker G., Volckaert G., Rombaux W., Billiau A., Somer P.;  
RT "Cloning of cDNA coding for human tissue-type plasminogen activator  
RT and its expression in *Escherichia coli*.";  
RL Mol. Biol. Med. 3:279-292(1986).  
RN [7]  
RP SEQUENCE FROM N.A. (ISOFORM SHORT).  
RC TISSUE=Umbilical vein;  
RA MEDLINE=90192129; PubMed=2107528;  
RA Siebert P.D., Fong K.;  
RT "Variant tissue-type plasminogen activator (PLAT) cDNA obtained from  
RT human endothelial cells.";  
RL Nucleic Acids Res. 18:1086-1086(1990).  
RN [8]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Brain;  
RA MEDLINE=22388257; PubMed=12477932;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raba S.S., Lequellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,  
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length  
RT human and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [9]  
RP SEQUENCE OF 212-361 FROM N.A.  
RA MEDLINE=83169656; PubMed=6572897;  
RA Edlund T., Ny T., Raanby M., Heden L.-O., Palm G., Holmgren E.,  
RA Josephson S.;  
RT "Isolation of cDNA sequences coding for a part of human tissue  
RT plasminogen activator.";  
RL Proc. Natl. Acad. Sci. U.S.A. 80:349-352(1983).  
RN [10]  
RP SEQUENCE OF 1-36 FROM N.A.  
RA MEDLINE=85289338; PubMed=3161893;  
RA Fisher R., Waller E.K., Grossi G., Thompson D., Tizard R.,  
RA Schleuning W.-D.;  
RT "Isolation and characterization of the human tissue-type plasminogen  
RT activator structural gene including its 5' flanking region.";  
RL J. Biol. Chem. 260:11223-11230(1985).  
RN [11]  
RP SEQUENCE OF 31-562 FROM N.A.  
RA MEDLINE=91291340; PubMed=1368681;  
RA Itagaki Y., Yasuda H., Morinaga T., Mitsuda S., Higashio K.;  
RT "Purification and characterization of tissue plasminogen activator  
RT secreted by human embryonic lung diploid fibroblasts, IMR-90 cells.";  
RL Agric. Biol. Chem. 55:1225-1232(1991).  
RN [12]  
RP SEQUENCE OF 36-562.  
RC TISSUE=Melanoma;  
RA MEDLINE=85000468; PubMed=6433976;  
RA Pohl G., Kaelstroom M., Bergsdorf N., Wallen P., Joernvall H.;  
RT "Tissue plasminogen activator: peptide analyses confirm an indirectly  
RT derived amino acid sequence, identify the active site serine residue,  
RT establish glycosylation sites, and localize variant differences.";  
RL Biochemistry 23:3701-3707(1984).  
RN [13]  
RP SEQUENCE OF 33-52 AND 311-330.  
RC TISSUE=Melanoma;  
RA MEDLINE=83209620; PubMed=6682760;  
RA Wallen P., Pohl G., Bergsdorf N., Raanby M., Ny T., Joernvall H.;  
RT "Purification and characterization of a melanoma cell plasminogen  
RT activator.";  
RL Eur. J. Biochem. 132:681-686(1983).  
RN [14]  
RP STRUCTURE OF CARBOHYDRATES.  
RA MEDLINE=90092112; PubMed=2513186;  
RA Pfeiffer G., Schmidt M., Strube K.-H., Geyer R.;  
RT "Carbohydrate structure of recombinant human uterine tissue  
RT plasminogen activator expressed in mouse epithelial cells.";  
RL Eur. J. Biochem. 186:273-286(1989).  
RN [15]  
RP CARBOHYDRATE-LINKAGE SITE THR-96.  
RA MEDLINE=91159408; PubMed=1900431;  
RA Harris R.J., Leonard C.K., Guzzetta A.W., Spellman M.W.;  
RT "Tissue plasminogen activator has an O-linked fucose attached to  
RT threonine-61 in the epidermal growth factor domain.";  
RL Biochemistry 30:2311-2314(1991).  
RN [16]  
RP DISULFIDE BONDS IN KRINGLE 2.  
RA MEDLINE=91244765; PubMed=1645336;  
RA Vlahos C.J., Wilhelm O.G., Hassell T., Jaskunas S.R., Bang N.U.;  
RT "Disulfide pairing of the recombinant kringle-2 domain of tissue  
RT plasminogen activator produced in *Escherichia coli*.";  
RL J. Biol. Chem. 266:10070-10072(1991).  
RN [17]  
RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF CATALYTIC DOMAIN.  
RA MEDLINE=96200985; PubMed=8613982;  
RA Lamba D., Bauer M., Huber R., Fischer S., Rudolph R., Kohnert U.,  
RA Bode W.;  
RT "The 2.3 A crystal structure of the catalytic domain of recombinant  
RT two-chain human tissue-type plasminogen activator.";  
RL J. Mol. Biol. 258:117-135(1996).  
RN [18]  
RP X-RAY CRYSTALLOGRAPHY (3.1 ANGSTROMS) OF CATALYTIC DOMAIN.  
RA MEDLINE=97449126; PubMed=9305622;  
RA Renatus M., Engh R.A., Stubbs M.T., Huber R., Fischer S., Kohnert U.,  
RA Bode W.;  
RT "Lysine 156 promotes the anomalous proenzyme activity of tPA: X-ray  
RT crystal structure of single-chain human tPA.";  
RL EMBO J. 16:4797-4805(1997).  
RN [19]  
RP X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS) OF KRINGLE 2.  
RA MEDLINE=92118803; PubMed=1310033;  
RA de Vos A., Ultsch M.H., Kelley R.F., Padmanabhan K., Tulinsky A.,  
RA Westbrook M.L., Kossiakof A.A.;  
RT "Crystal structure of the kringle 2 domain of tissue plasminogen  
RT activator at 2.4-A resolution.";  
RL Biochemistry 31:270-279(1992).  
RN [20]  
RP STRUCTURE BY NMR OF KRINGLE 2.  
RA MEDLINE=90122799; PubMed=2558718;  
RA Byeon I.-J.L., Kelley R.F., Llinas M.;  
RT "1H NMR structural characterization of a recombinant kringle 2 domain  
RT from human tissue-type plasminogen activator.";  
RL Biochemistry 28:9350-9360(1989).  
RN [21]  
RP STRUCTURE BY NMR OF KRINGLE 2.  
RA MEDLINE=91200042; PubMed=1901789;  
RA Byeon I.-J.L., Kelley R.F., Llinas M.;  
RT "Kringle-2 domain of the tissue-type plasminogen activator. 1H-NMR  
RT assignments and secondary structure.";  
RL Eur. J. Biochem. 197:155-165(1991).  
RN [22]  
RP STRUCTURE BY NMR OF KRINGLE 2.  
RA MEDLINE=92106329; PubMed=1762144;  
RA Byeon I.-J.L., Llinas M.;  
RT "Solution structure of the tissue-type plasminogen activator kringle  
RT 2 domain complexed to 6-aminohexanoic acid an antifibrinolytic  
RT drug.";  
RL J. Mol. Biol. 222:1035-1051(1991).  
RN [23]



Query Match 95.1%; Score 1961; DB 1; Length 562;  
Best Local Similarity 99.7%; Pred. No. 1.1e-162;  
Matches 355; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 21 AASEGSDCVPGNSAYRGTHSLTSGASCLPWNMILIGKVTYTAQNPASQAALGLGKHNY 80  
DB 207 ACSENGSDCVPGNSAYRGTHSLTSGASCLPWNMILIGKVTYTAQNPASQAALGLGKHNY 266

QY 81 CRNPDGDAKWPCHVLYKRNRLTWECVPSCTGLQYVSPQFRIKGLFADIASHPWQA 140  
DB 267 CRNPDGDAKWPCHVLYKRNRLTWECVPSCTGLQYVSPQFRIKGLFADIASHPWQA 326

QY 141 AIPAKHRRSPGERFLCGILLISSCWILSAHCFQERPPPHHLTVILGRTVYVPGEBEQK 200  
DB 327 AIPAKHRRSPGERFLCGILLISSCWILSAHCFQERPPPHHLTVILGRTVYVPGEBEQK 386

QY 201 FEVEKYIVHKEFDVTDNDIALQKSDSRCAQESSVVRTVCLPPADLQLPDWTCECL 260  
DB 387 FEVEKYIVHKEFDVTDNDIALQKSDSRCAQESSVVRTVCLPPADLQLPDWTCECL 446

QY 261 SGYKHEALSPFYSERLKEARVLYPSRCTSQHLLNRTVTDNMLCAGDTRSGGQANLH 320  
DB 447 SGYKHEALSPFYSERLKEARVLYPSRCTSQHLLNRTVTDNMLCAGDTRSGGQANLH 506

QY 321 DACQDGGPLVCLNDGRMTLVGIISWGLGCGKQDVGVTYKVTNYLDWIRDNNRP 376  
DB 507 DACQDGGPLVCLNDGRMTLVGIISWGLGCGKQDVGVTYKVTNYLDWIRDNNRP 562

RESULT 2

TPA\_BOVIN STANDARD; PRT; 566 AA.

AC Q28198;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Tissue-type plasminogen activator precursor (EC 3.4.21.68) (tPA)  
DE (t-PA) (t-plasminogen activator).  
GN PLAT.  
OS Bos taurus (Bovine).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
OC Bovidae; Bovinae; Bos.  
OX NCBI\_TaxID=9913;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE-KIDNEY;  
RA Ravn P., Berglund L., Petersen T.E.;  
RT "Cloning and characterization of the bovine plasminogen activators uPA  
and tPA.";  
RL Int. Dairy J. 5:605-617(1995).  
CC -1- FUNCTION: CONVERTS THE ABUNDANT, BUT INACTIVE, ZYMOMEN PLASMINOGEN  
TO PLASMIN BY HYDROLYZING A SINGLE ARG-VAL BOND IN PLASMINOGEN. BY  
CONTROLLING PLASMIN-MEDIATED PROTEOLYSIS, IT PLAYS AN IMPORTANT  
ROLE IN TISSUE REMODELING AND DEGRADATION, IN CELL MIGRATION AND  
MANY OTHER PHYSIOLOGICAL EVENTS.  
CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in  
plasminogen to form plasmin.  
CC -1- SUBUNIT: HETERODIMER OF CHAIN A AND CHAIN B HELD BY A DISULFIDE  
BOND.  
CC -1- SUBCELLULAR LOCATION: SECRETED; EXTRACELLULAR.  
CC -1- PTM: THE SINGLE CHAIN, ALMOST FULLY ACTIVE ENZYME, CAN BE FURTHER  
PROCESSED INTO A TWO-CHAIN FULLY ACTIVE FORM BY A CLEAVAGE AFTER  
ARG-314 CATALYZED BY PLASMIN, TISSUE KALLIKREIN OR FACTOR XA.  
CC -1- MISCELLANEOUS: BINDS TO THE KRINGLE STRUCTURE OF THE FIBRIN A  
CHAIN. BINDING TO FIBRIN ENHANCES ITS CATALYTIC ACTIVITY.  
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.  
CC -1- SIMILARITY: Contains 1 EGF-like domain.  
CC -1- SIMILARITY: Contains 1 fibronectin type I domain.  
CC -1- SIMILARITY: Contains 2 kringle domains.

-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (see <http://www.isb-sib.ch/announce/>)  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
CC EMBL; X85800; CAA59795.1; -;  
CC HSP; P00750; IRTF.  
CC MEROPS; S01.232; -;  
CC InterPro; IPR001314; Chymotrypsin.  
CC InterPro; IPR006209; EGF\_like.  
CC InterPro; IPR000083; Fibrinctnl.  
CC InterPro; IPR006210; IEGF.  
CC InterPro; IPR000001; Kringle.  
CC InterPro; IPR01254; Ser\_protease\_Try.  
CC Pfam; PF00008; EGF; 1.  
CC Pfam; PF00039; fn1; 1.  
CC Pfam; PF00051; kringle; 2.  
CC Pfam; PF00089; trypsin; 1.  
CC PRINTS; PR00722; CHYMOTRYPSIN.  
CC PRINTS; PR0018; KRINGLE.  
CC ProDom; PD000395; Kringle; 2.  
CC SMART; SM00181; EGF; 1.  
CC SMART; SM00058; FN1; 1.  
CC SMART; SM00130; KR; 2.  
CC SMART; SM00020; Tryp\_Spc; 1.  
CC PROSITE; PS00022; EGF\_1; 1.  
CC PROSITE; PS01186; EGF\_2; 1.  
CC PROSITE; PS01253; FIBRONECTIN\_1; 1.  
CC PROSITE; PS00021; KRINGLE\_1; 1.  
CC PROSITE; PS00070; KRINGLE\_2; 2.  
CC PROSITE; PS00240; TRYPSIN\_DOM; 1.  
CC PROSITE; PS00134; TRYPSIN\_HIS; 1.  
CC PROSITE; PS00135; TRYPSIN\_SER; 1.  
KW Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;  
FT SIGNAL 1 21 BY SIMILARITY.  
FT PROPEP 22 33 BY SIMILARITY.  
FT CHAIN 34 566 TISSUE-TYPE PLASMINOGEN ACTIVATOR A  
FT CHAIN 34 314 TISSUE-TYPE PLASMINOGEN ACTIVATOR B  
FT CHAIN 315 566 CHAIN.  
FT DOMAIN 40 82 FIBRONECTIN TYPE-I.  
FT DOMAIN 83 121 EGF-LIKE.  
FT DOMAIN 128 209 KRINGLE 1.  
FT DOMAIN 219 300 KRINGLE 2.  
FT DOMAIN 315 566 SERINE PROTEASE.  
FT ACT\_SITE 361 361 CHARGE RELAY SYSTEM.  
FT ACT\_SITE 410 410 CHARGE RELAY SYSTEM.  
FT ACT\_SITE 517 517 CHARGE RELAY SYSTEM.  
FT DISULFID 42 72 BY SIMILARITY.  
FT DISULFID 70 79 BY SIMILARITY.  
FT DISULFID 87 98 BY SIMILARITY.  
FT DISULFID 92 109 BY SIMILARITY.  
FT DISULFID 111 120 BY SIMILARITY.  
FT DISULFID 128 209 BY SIMILARITY.  
FT DISULFID 149 191 BY SIMILARITY.  
FT DISULFID 180 204 BY SIMILARITY.  
FT DISULFID 219 300 BY SIMILARITY.  
FT DISULFID 240 282 BY SIMILARITY.  
FT DISULFID 271 295 BY SIMILARITY.  
FT DISULFID 303 434 INTERCHAIN (BY SIMILARITY).  
FT DISULFID 346 362 BY SIMILARITY.  
FT DISULFID 354 423 BY SIMILARITY.  
FT DISULFID 448 523 BY SIMILARITY.  
FT DISULFID 480 496 BY SIMILARITY.  
FT DISULFID 513 541 BY SIMILARITY.  
FT CARBOHYD 153 153 N-LINKED (GLCNAC... ) (POTENTIAL).  
FT CARBOHYD 487 487 N-LINKED (GLCNAC... ) (POTENTIAL).  
SQ SEQUENCE 566 AA; 63701 MW; 2EB6BEB4E32276C3 CRC64;



```

FT DISULFID 38 68 BY SIMILARITY.
FT DISULFID 66 75 BY SIMILARITY.
FT DISULFID 83 94 BY SIMILARITY.
FT DISULFID 88 105 BY SIMILARITY.
FT DISULFID 107 116 BY SIMILARITY.
FT DISULFID 124 205 BY SIMILARITY.
FT DISULFID 145 187 BY SIMILARITY.
FT DISULFID 176 200 BY SIMILARITY.
FT DISULFID 213 294 BY SIMILARITY.
FT DISULFID 234 276 BY SIMILARITY.
FT DISULFID 265 289 BY SIMILARITY.
FT DISULFID 297 428 INTERCHAIN (BY SIMILARITY).
FT DISULFID 340 356 BY SIMILARITY.
FT DISULFID 348 417 BY SIMILARITY.
FT DISULFID 442 516 BY SIMILARITY.
FT DISULFID 474 490 BY SIMILARITY.
FT DISULFID 506 534 BY SIMILARITY.
FT CARBOHYD 149 149 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 481 481 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 380 380 E -> K (IN REF. 1).
SQ SEQUENCE 559 AA; 62903 MW; 7DBD3809C1D1C921 CRC64;

Query Match 79.5%; Score 1638.5; DB 1; Length 559;
Best Local Similarity 82.5%; Pred No. 1.1e-134;
Matches 287; Conservative 29; Mismatches 31; Indels 1; Gaps 1;

QY 28 DCYFGNGSAYRGTHSLTSGASCLPWNMSMILIGKVTYTAQNPASQAQGLGKHNCRNPDGD 87
DB 212 DCYVGKGYVTRGTHSTFTSKASCLPWNMSMILIGKVTYTAQNPASQAQGLGKHNCRNPDGD 271

QY 88 AKPCHVLKRNRLTWEYCDVSCSTGCLRQYQSOPQFRIGKGLPADIASHPWQAIFAKHR 147
DB 272 AKPCHVNMKRLTWEYCDVSCSTGCLRQYQSOPQFRIGKGLPADIASHPWQAIFAKHR 331

QY 148 RSPGRFLCGILISSCWILSAACHFCQRRPPHLLTVILGRTYRVVPGEEBQKEVEKYI 207
DB 332 RSPGRFLCGVGLISSCWILSAACHFCQRRPPHLLTVILGRTYRVVPGEEBQTEIEKYI 391

QY 208 VHKFDDDDYNDIALQLKSDSRCAQESSVWRTVCLPPADLQLPDWTCELSGYGKHE 267
DB 392 VHKFDDDDYNDIALQLKSDSRCAQESSVWRTVCLPPADLQLPDWTCELSGYGKHE 451

QY 268 ALSFPYSRLKEAHVRLYPSRRTSOHLNLTNTVDMNCAGTSGGSGPQANLHACQGD 327
DB 452 ASSPFFSRLKEAHVRLYPSRRTSOHLNLTNTVDMNCAGTSGGSGPQANLHACQGD 510

QY 328 GGPLVCLNDGRMTLVGLISWGLGCGKQDVPVGVYTKVTNLYLDWIRDMNR 375
DB 511 GGPLVCLNDGRMTLVGLISWGLGCGKQDVPVGVYTKVTNLYLDWIRDMNR 558

RESULT 4
TPA_MOUSE STANDARD; PRT; 559 AA.
AC P11214; Q91VP2;
DT 01-JUL-1989 (Rel. 11, Created)
DT 15-SEP-2003 (Rel. 42, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Tissue-type plasminogen activator precursor (EC 3.4.21.68) (tPA)
DE (t-PA) (t-plasminogen activator).
GN PLAT.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88087303; PubMed=2826484;
RA Rickles R.J., Darlow A.L., Strickland S.;
RT "Molecular cloning of complementary DNA to mouse tissue plasminogen
RT activator mRNA and its expression during F9 teratocarcinoma cell
RT differentiation.";
RL J. Biol. Chem. 263:1563-1569(1988).

```

```

RN RP SEQUENCE FROM N.A.
RC TISSUE=Mammary gland;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heide F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Tohiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalek U., Smallos D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -!- FUNCTION: CONVERTS THE ABUNDANT, BUT INACTIVE, ZIMOGEN PLASMINOGEN
CC TO PLASMIN BY HYDROLYZING A SINGLE ARG-VAL BOND IN PLASMINOGEN. BY
CC CONTROLLING PLASMIN-MEDIATED PROTEOLYSIS, IT PLAYS AN IMPORTANT
CC ROLE IN TISSUE REMODELING AND DEGRADATION, IN CELL MIGRATION AND
CC MANY OTHER PHYSIOLOGICAL EVENTS.
CC -!- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
CC plasminogen to form plasmin.
CC -!- SUBUNIT: HETERODIMER OF CHAIN A AND CHAIN B HELD BY A DISULFIDE
CC BOND.
CC -!- SUBCELLULAR LOCATION: SECRETED; EXTRACELLULAR.
CC -!- PTM: THE SINGLE CHAIN, ALMOST FULLY ACTIVE ENZYME, CAN BE FURTHER
CC PROCESSED INTO A TWO-CHAIN FULLY ACTIVE FORM BY A CLEAVAGE AFTER
CC ARG-308 CATALYZED BY PLASMIN, TISSUE KALLIKREIN OR FACTOR XA.
CC -!- MISCELLANEOUS: BINDS TO THE KRINGLE STRUCTURE OF THE FIBRIN A
CC CHAIN. BINDING TO FIBRIN ENHANCES ITS CATALYTIC ACTIVITY.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -!- SIMILARITY: Contains 1 EGF-like domain.
CC -!- SIMILARITY: Contains 1 fibronectin type I domain.
CC -!- SIMILARITY: Contains 2 kringle domains.
CC -!- SIMILARITY: Contains 2 kringle domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; J03520; AAA0470.1; --
CC EMBL; BC011256; AAH11256.1; --
CC PIR; A29941; A29941.
CC HSHP; P00750; IA5H.
CC MEROPS; S01.232; --
CC MGD; MGI-97610; Plat.
CC InterPro; IPR001314; Chymotrypsin.
CC InterPro; IPR006209; EGF_like.
CC InterPro; IPR000083; Fibnctnl.
CC InterPro; IPR006210; IEGF.
CC InterPro; IPR000001; Kringle.
CC InterPro; IPR001254; Ser_protease_Try.
CC Pfam; PF00008; EGF; 1.
CC Pfam; PF00039; fn1; 1.
CC Pfam; PF00051; kringle; 2.
CC Pfam; PF00089; trypsin; 1.
CC PRINTS; PR00722; CHYMOTRYPSIN.
CC PRINTS; PR00018; KRINGLE.
CC ProDom; PD000395; Kringle; 2.
CC SMART; SM00181; EGF; 1.
CC SMART; SM00058; FN1; 1.

```

DR SMART; SM00130; KR; 2.  
DR SMART; SM00020; Tryp SPC; 1.  
DR PROSITE; PS00022; EGF\_1; 1.  
DR PROSITE; PS01186; EGF\_2; 1.  
DR PROSITE; PS01253; FIBRONECTIN\_1; 1.  
DR PROSITE; PS00021; KRINGLE\_1; 2.  
DR PROSITE; PS00070; KRINGLE\_2; 2.  
DR PROSITE; PS00240; TRYPSIN\_DOM; 1.  
DR PROSITE; PS00134; TRYPSIN\_HIS; 1.  
DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
KW Plasminogen activator; Hydrolase; Serine protease; Glycoprotein;  
KW Plasma; Kringle; EGF-like domain; Repeat; Signal.  
FT SIGNAL 1 17  
FT PROPEP 18 29  
FT CHAIN 30 559  
FT CHAIN 30 308  
FT CHAIN 309 559  
FT DOMAIN 36 78  
FT DOMAIN 79 117  
FT DOMAIN 124 205  
FT DOMAIN 213 294  
FT DOMAIN 309 559  
FT ACT SITE 355 355  
FT ACT SITE 404 404  
FT ACT SITE 510 510  
FT DISULFID 38 68  
FT DISULFID 66 75  
FT DISULFID 83 94  
FT DISULFID 88 105  
FT DISULFID 107 116  
FT DISULFID 124 205  
FT DISULFID 145 187  
FT DISULFID 176 200  
FT DISULFID 213 294  
FT DISULFID 234 276  
FT DISULFID 265 289  
FT DISULFID 297 428  
FT DISULFID 340 356  
FT DISULFID 348 417  
FT DISULFID 442 516  
FT DISULFID 474 490  
FT DISULFID 506 534  
FT CARBOHYD 149 149  
FT CARBOHYD 481 481  
FT CONFLICT 260 260  
FT CONFLICT 325 325  
SQ SEQUENCE 559 AA; 63122 MW; 8CCE2BDB94514D9 CRC64;  
  
Query Match  
Best Local Similarity 77.7%; Score 1601; DB 1; Length 559;  
Matches 283; Conservative 34; Mismatches 37; Indels 2; Gaps 2;  
  
QY 21 AASEGNS-DCYFGNSAYRGTHSLTSGASCLPWNMSILIGKVTYTAQNPQAALGLGKH 79  
DB 204 ACPKGSDECYGVGTYRGTHSLTTSQASCLPWNMSIVLMGKSYTAWNTNSQALGLGHN 263  
  
QY 80 YCRNPDGAKPWCHVKLRRLTWECYDVPSCSTGLROYSPQPRFKGLFADTASHPWQ 139  
DB 264 YCRNPDGAPWCHVKMDKRLTWECYDMSPCSTGLROYKRPQPRFKGLYDITSHPWQ 323  
  
QY 140 AAIFAKHRRSPGERFLCGGILISSCWILSAACHFCRPPPHLTVILGRTVRVVPGEEQ 199  
DB 324 APIFVKNRSPGERFLCGGVLISSCWLSAAHCFERPPPHLKVILGRTVRVVPGEEQ 383  
  
QY 200 KFEYKIVHKEFDYDDTDYNDIALQLKSDSRCAQESSVVRTVCLPPADLQLPDWTECE 259  
DB 384 TFEIKYIVHEEFDYDDTDYNDIALQLKSSQKCAQESSVGTACLPPDNLQLPDWTECE 443  
  
QY 260 LSGYKHEALSPFFYSERLKEAHRVLPSSRCTSQHLNRTVDNMLCAGDTRSGGQANL 319  
DB 444 LSGYKHEASPPFFSDRLKEAHRVLPSSRCTSQHLFNKNTNNMLCAGDTRSGGQ-DL 502

QY 320 HDACQGDSDGGLVCLANDGRMTLVGIISWGLGCGKQVPGVYTKVTNYLDWIRNMR 375  
DB 503 HDACQGDSDGGLVCLANDGRMTLVGIISWGLGCGKQVPGVYTKVTNYLDWIRNMR 558

## RESULT 5

URTG DESRO  
ID URTG DESRO STANDARD; PRT; 394 AA.  
AC P49150;  
DT 01-FEB-1996 (Rel. 33, Created)  
DT 01-FEB-1996 (Rel. 33, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Salivary plasminogen activator gamma precursor (EC 3.4.21.68) (DSPA gamma).  
DE Desmodus rotundus (Vampire bat).  
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Chiroptera; Microchiroptera; Phyllostomidae;  
OC Desmodontinae; Desmodus.  
OX NCBI\_TaxID=9430;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Salivary gland;  
RX MEDLINE=92039036; PubMed=1937019;  
RA Kraetzschmar J., Haendler B., Langer G., Boidol W., Bringmann P.,  
RA Alagon A., Donner P., Schleuning W.D.;  
RT "The plasminogen activator family from the salivary gland of the  
vampire bat Desmodus rotundus: cloning and expression.";  
RL Gene 105:229-237(1991).  
RN [2]  
RP CHARACTERIZATION.  
RX MEDLINE=93393059; PubMed=1309059;  
RA Schleuning W.-D., Alagon A., Boidol W., Bringmann P., Petri T.,  
RA Kraetzschmar J., Haendler B., Langer G., Baldus B., Witt W.,  
RA Donner P.;  
RT "Plasminogen activators from the saliva of Desmodus rotundus (common  
vampire bat): unique fibrin specificity.";  
RL Ann. N.Y. Acad. Sci. 667:395-403(1992).  
CC -!- FUNCTION: PROBABLY ESSENTIAL TO SUPPORT THE FEEDING HABITS OF THIS  
EXCLUSIVELY HAEMATOPHAGOUS ANIMAL. PROBABLE POTENT THROMBOLYTIC  
AGENT.  
CC -!- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in  
plasminogen to form plasmin.  
CC -!- SUBUNIT: Monomer.  
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.  
CC -!- SIMILARITY: Contains 1 kringle domain.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; M63990; AAA31595.1; -.  
DR PIR; J06600; J06600.  
DR HSP; P98119; IAS1.  
DR MEROPS; S01.239; -.  
DR InterPro; IPR001314; Chymotrypsin.  
DR InterPro; IPR000001; Kringle.  
DR InterPro; IPR001254; Ser. protease\_Try.  
DR Pfam; PF00051; kringle\_1.  
DR Pfam; PF00089; trypsin\_1.  
DR PRINTS; PR00722; CHYMOTRYPSIN.  
DR PRINTS; PR00018; KRINGLE.  
DR ProDom; PD000395; Kringle; 1.  
DR SMART; SM00130; KR; 1.  
DR SMART; SM00020; Tryp SPC; 1.  
DR PROSITE; PS00021; KRINGLE\_1; 1.  
DR PROSITE; PS00070; KRINGLE\_2; 1.  
DR PROSITE; PS00240; TRYPSIN\_DOM; 1.  
DR PROSITE; PS00134; TRYPSIN\_HIS; 1.  
DR PROSITE; PS00135; TRYPSIN\_SER; 1.

Kraetzschmar J., Haendler B., Langer G., Boidol W., Bringmann P., Alagon A., Donner P., Schleuning W.D.;  
 "The plasminogen activator family from the salivary gland of the vampire bat *Desmodus rotundus*: cloning and expression.";  
 Gene 105:229-237(1991).  
 [2]  
 RN SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
 RP TISSUE=Salivary gland;  
 RX MEDLINE=9003687; PubMed=2509450;  
 RA Gardell S.J., Duong L.T., Diehl R.E., York J.D., Hare T.R., Register R.B., Jacobs J.W., Dixon R.A.F., Friedman P.A.;  
 "Isolation, characterization, and cDNA cloning of a vampire bat salivary plasminogen activator.";  
 J. Biol. Chem. 264:17947-17952(1989).  
 [3]  
 RN CHARACTERIZATION.  
 RP MEDLINE=93393059; PubMed=1305059;  
 RX Schleuning W.-D., Alagon A., Boidol W., Bringmann P., Petri T., Kraetzschmar J., Haendler B., Langer G., Baldus B., Witt W., Donner P.;  
 "Plasminogen activators from the saliva of *Desmodus rotundus* (common vampire bat): unique fibrin specificity.";  
 Ann. N.Y. Acad. Sci. 667:395-403(1992).  
 CC -1- FUNCTION: PROBABLY ESSENTIAL TO SUPPORT THE FEEDING HABITS OF THIS EXCLUSIVELY HAEMATOPHAGOUS ANIMAL. PROBABLE POTENT THROMBOLYTIC AGENT.  
 CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in plasminogen to form plasmin.  
 CC -1- ENZYME REGULATION: ACTIVITY TOWARD PLASMINOGEN IS STIMULATED IN THE PRESENCE OF FIBRIN I.  
 CC -1- SUBUNIT: Monomer.  
 CC -1- DOMAIN: THE FIBRONECTIN TYPE-I DOMAIN MEDIATES BINDING TO FIBRIN, AND THE KRINGLE DOMAIN APPARENTLY MEDIATES FIBRIN-INDUCED STIMULATION OF ACTIVITY.  
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.  
 CC -1- SIMILARITY: Contains 1 egf-like domain.  
 CC -1- SIMILARITY: Contains 1 fibronectin type I domain.  
 CC -1- SIMILARITY: Contains 1 kringle domain.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announcement/>) or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----  
 DR EMBL; M63988; AAA31593.1; -;  
 DR EMBL; J05082; AAA31596.1; -;  
 DR PIR; J05098; J05098.  
 DR HSSP; P98119; 1A51.  
 DR MEROPS; S01.232; -;  
 DR InterPro; IPR001314; Chymotrypsin.  
 DR InterPro; IPR006209; EGF\_like.  
 DR InterPro; IPR000083; Fibnctnl.  
 DR InterPro; IPR006210; IEGF.  
 DR InterPro; IPR000001; Kringle.  
 DR InterPro; IPR001254; Ser\_protease\_Try.  
 DR Pfam; PF00008; EGF; 1.  
 DR Pfam; PF00039; fni; 1.  
 DR Pfam; PF00051; kringle; 1.  
 DR Pfam; PF00089; trypsin; 1.  
 DR PRINTS; PR00722; CHYMOTRYPSIN.  
 DR PRINTS; PR00018; KRINGLE.  
 DR ProDom; PD000395; Kringle; 1.  
 DR SMART; SM00181; EGF; 1.  
 DR SMART; SM00058; FN1; 1.  
 DR SMART; SM00130; KR; 1.  
 DR SMART; SM00020; Tryp\_Spc; 1.  
 DR PROSITE; PS00022; EGF\_1; 1.  
 DR PROSITE; PS01186; EGF\_2; 1.  
 DR PROSITE; PS01253; FIBRONECTIN\_1; 1.  
 DR PROSITE; PS00021; KRINGLE\_1; 1.

DR PROSITE; PS00070; KRINGLE 2; 1.  
 DR PROSITE; PS00240; TRYPSIN\_DOM; 1.  
 DR PROSITE; PS00134; TRYPSIN\_HIS; 1.  
 DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
 KW Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;  
 KRingle; EGF-like domain; Signal; Multigene family.  
 FT SIGNAL 1 36 POTENTIAL.  
 FT CHAIN 37 477 SALIVARY PLASMINOGEN ACTIVATOR ALPHA 2.  
 FT DOMAIN 40 82 FIBRONECTIN TYPE-I.  
 FT DOMAIN 83 121 EGF-LIKE.  
 FT DOMAIN 128 209 KRINGLE.  
 FT DOMAIN 225 477 SERINE PROTEASE.  
 FT ACT\_SITE 272 477 CHARGE RELAY SYSTEM (BY SIMILARITY).  
 FT ACT\_SITE 321 428 CHARGE RELAY SYSTEM (BY SIMILARITY).  
 FT ACT\_SITE 428 428 CHARGE RELAY SYSTEM (BY SIMILARITY).  
 FT DISULFID 42 72 BY SIMILARITY.  
 FT DISULFID 70 79 BY SIMILARITY.  
 FT DISULFID 87 98 BY SIMILARITY.  
 FT DISULFID 92 109 BY SIMILARITY.  
 FT DISULFID 111 120 BY SIMILARITY.  
 FT DISULFID 128 209 BY SIMILARITY.  
 FT DISULFID 149 191 BY SIMILARITY.  
 FT DISULFID 180 204 BY SIMILARITY.  
 FT DISULFID 214 345 BY SIMILARITY.  
 FT DISULFID 257 273 BY SIMILARITY.  
 FT DISULFID 265 334 BY SIMILARITY.  
 FT DISULFID 359 434 BY SIMILARITY.  
 FT DISULFID 391 407 BY SIMILARITY.  
 FT DISULFID 424 452 BY SIMILARITY.  
 FT CARBOHYD 185 185 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 398 398 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CONFLICT 403 403 N -> K (IN REF. 2).  
 FT CONFLICT 417 417 Y -> H (IN REF. 2).  
 FT CONFLICT 435 435 M -> R (IN REF. 2).  
 SQ SEQUENCE 477 AA; 53719 MW; 17486555C0E3077C CRC64;

Query Match 69.3%; Score 1428; DB 1; Length 477;  
 Best Local Similarity 70.2%; Pred. No. 1.8e-116;  
 Matches 254; Conservative 40; Mismatches 66; Indels 2; Gaps 1;

QY 17 TVQAASEGSDCYFGNGSAVRGTHSLTESGASCLPWNMILGKVTYTAQNPQAALGLG 76  
 DB 116 TGKQCEVDTHATCYKQDQGVYRGTWSTSESQAQINMNSNLLTRTYNGRSDAITLGLG 175  
 QY 77 KHNCRNPDGAKPWCHVLKRNRLTWECYDVPSCS--TCGLRQYSDQFRKGLGFADIA 134  
 DB 176 NHNCRNPDNNKPCWYVVKASKFILEFCFVPVCSKATCGLRKYKBPQLHSTGGLFTDIT 235  
 QY 135 SHPWQAALFAKHRSFGRFLCGGILISSCWILSAAHCFQERFPFPHLTIVILGRTYRVVP 194  
 DB 236 SHPWQAALFAQNRSSGRFLCGGILISSCWILTAHCFQERYPPQHLRVVLGRTYRVKP 295  
 QY 195 GEESQKFEVKYIVHKEPDDTYNDIALQLKSDSRCAQESSVVRTVCLPPADLQLPD 254  
 DB 296 GKKEQTEVEKICVHVEEFDYNDIALLQLKSGSPQCAQESDVSRAICLPEANLQLPD 355  
 QY 255 WTECELSGVGKHEALSPPYSRLKAHVRLYPSSRCTSOHLNATVTNDMLCAGDTSSG 314  
 DB 356 WTECELSGVGKHKSSPPYSQLEKGVRLYPSSRCTSKFLNKVTNNMLCAGDTSSGE 415  
 QY 315 PQANLHADACQDGGPLVCLNDGRMTLVGIISWGLGCGQKQDVPGVYTKVTNYLDWIRDNM 374  
 DB 416 IYPNVHADACQDGGPLVCMNDNHTLLGIISWVGCGEKDIPGVYTKVTNYLGIWIRDNM 475  
 QY 375 RP 376  
 DB 476 RP 477

RESULT 7

URTB\_DESRO

ID URTB\_DESRO

AC P98121;

STANDARD;

PRT;

431 AA.

DT 01-FEB-1996 (Rel. 33, Created)  
 DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Salivary plasminogen activator beta precursor (EC 3.4.21.68) (DSPA beta).  
 DE Desmodus rotundus (Vampire bat).  
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Chiroptera; Microchiroptera; Phyllostomidae; Desmodontinae; Desmodus.  
 OC NCBI\_TaxID=9430;  
 OX [1]  
 RN SEQUENCE FROM N.A.  
 RP TISSUE=Salivary gland;  
 RC MEDLINE=92039036; PubMed=1937019;  
 RX Kraetzschmar J., Haendler B., Langer G., Boidol W., Bringmann P., Alagon A., Donner P., Schleuning W.D.;  
 RA "The plasminogen activator family from the salivary gland of the vampire bat Desmodus rotundus: cloning and expression.";  
 RT Gene 105:229-237(1991).  
 RL [2]  
 RN CHARACTERIZATION.  
 RP MEDLINE=9339059; PubMed=1309059;  
 RA Schleuning W.-D., Alagon A., Boidol W., Bringmann P., Petri T., Kraetzschmar J., Haendler B., Langer G., Baldus B., Witt W., Donner P.;  
 RA "Plasminogen activators from the saliva of Desmodus rotundus (common vampire bat): unique fibrin specificity.";  
 RT Ann. N.Y. Acad. Sci. 667:395-403(1992).  
 RL CC -1- FUNCTION: PROBABLY ESSENTIAL TO SUPPORT THE FEEDING HABITS OF THIS EXCLUSIVELY HAEMATOPHAGOUS ANIMAL. PROBABLE POTENT THROMBOLYTIC AGENT.  
 CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in plasminogen to form plasmin.  
 CC -1- SUBUNIT: Monomer.  
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.  
 CC -1- SIMILARITY: Contains 1 kringle domain.  
 CC -1- SIMILARITY: Contains 1 EGF-like domain.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; M63989; AAA31594.1; -.  
 DR PIR; JS0599; JS0599.  
 DR HSPP; P98119; IA51.  
 DR MEROPS; S01.239; -.  
 DR InterPro; IPR001314; Chymotrypsin.  
 DR InterPro; IPR006209; EGF\_like.  
 DR InterPro; IPR006210; IEGF.  
 DR InterPro; IPR000001; Kringle.  
 DR InterPro; IPR001254; Ser\_protease\_Try.  
 DR Pfam; PF00008; EGF; 1.  
 DR Pfam; PF00051; kringle; 1.  
 DR Pfam; PF00089; trypsin; 1.  
 DR PRINTS; PR00722; CHYMOTRYPSIN.  
 DR PRINTS; PR00018; KRINGLE.  
 DR PRODOM; PD000395; Kringle; 1.  
 DR SMART; SM00181; EGF; 1.  
 DR SMART; SM00130; KR; 1.  
 DR SMART; SM00020; Tryp\_SPC; 1.  
 DR PROSITE; PS00022; EGF\_1; 1.  
 DR PROSITE; PS01186; EGF\_2; 1.  
 DR PROSITE; PS00021; KRINGLE\_1; 1.  
 DR PROSITE; PS00070; KRINGLE\_2; 1.  
 DR PROSITE; PS50070; TRYPSIN\_DOM; 1.  
 DR PROSITE; PS00240; TRYPSIN\_HIS; 1.  
 DR PROSITE; PS00134; TRYPSIN\_SER; 1.  
 DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
 KW Plasminogen activation; Hydrolase; Serine protease; Glycoprotein; Kringle; EGF-like domain; Signal; Multigene family.

FT SIGNAL 1 36  
FT CHAIN 37 431  
FT DOMAIN 37 75  
FT DOMAIN 82 163  
FT DOMAIN 179 431  
FT ACT\_SITE 226 226  
FT ACT\_SITE 275 275  
FT ACT\_SITE 382 382  
FT DISULFID 41 52  
FT DISULFID 46 63  
FT DISULFID 65 74  
FT DISULFID 82 163  
FT DISULFID 103 145  
FT DISULFID 134 158  
FT DISULFID 168 299  
FT DISULFID 211 227  
FT DISULFID 219 288  
FT DISULFID 313 388  
FT DISULFID 345 361  
FT DISULFID 378 406  
FT CARBOHYD 139 139  
FT CARBOHYD 352 352  
SQ SEQUENCE 431 AA; 48221 MW; 699B5B675B162CBF CRC64;

Query Match 69.2%; Score 1427; DB 1; Length 431;  
Best Local Similarity 69.9%; Pred. No. 1.9e-116;  
Matches 253; Conservative 41; Mismatches 66; Indels 2; Gaps 1;

17 TVQAASEGSDCYFGNGSAVRGTHSLTESGASCLPWNMSLLIGKVTYTAQNPSSAQAALGLG 76  
70 TGKQCEVDTHATCYKQGVYTRGTWSTSGAQCNWNSLLTRRTYNGRRSDAITLGLG 129  
77 KHNYCRNPDGAKPWCHLVKNRLTWECYDVPCS--TCGLRQYSQPFRIKGLGFADIA 134  
130 NHNYCRNPDNNSKPCVYKASKFILEPCSVPCSKATCGLRKYKEQLHSTGGLFTDIT 189  
135 SHPQQAIFAKHRSRPERFLCGGILSSCWILSAACHCFQRPFPHPHLLTVILGTYRVVP 194  
190 SHPQQAIFAQNRSSGERFLCGGILSSCWLVLTAAHCFQRYPPQHLRVVLGRTYRVKP 249  
195 GEESQKEVEKYIVHKEFDDTDYNDIALLOKSDSRCAQESSVVRVTCVLPADLOLPD 254  
250 GKESQTEVEKCIIEHFDDTDYNDIALLOKSGSPQCAQESSVRAICLPEANLQLPD 309  
255 WTECELSGYGKHGHEALSPFYSERLKEAHVRLYPSSRCTSQHLNRTVTDNMLCAGDTRSGG 314  
310 WTECELSGYGKHGSSPPFYSEQLKEGHVRLYPSSRCTSKFLFNKVTYTNMLCAGDTRSGE 369  
315 PQANLHDACQDGGPLVCLNDGRMTLVGIISWGLCGCKQKDPVGYTKVTNYLDWIRDNM 374  
370 IYPNVHDACQDGGPLVCMNDNHTLLGIISWGGCGEKDIPGVYTKVTNYLGLWIRDNM 429  
375 RP 376  
430 RP 431

STANDARD; PRT; 477 AA.

RESULT 8  
ID \_URL1 DESRO  
AC P98119;  
DT 01-FEB-1996 (Rel. 33, Created)  
DT 01-FEB-1996 (Rel. 33, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Salivary plasminogen activator alpha 1 precursor (EC 3.4.21.68) (DSFA alpha-1).  
OS Desmodus rotundus (Vampire bat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Chiroptera; Microchiroptera; Phyllostomidae;  
OC Desmodontinae; Desmodus.  
OX NCBI\_TaxID=9430;  
RN [1]  
RP SEQUENCE FROM N.A.

RC TISSUE=Salivary gland;  
RX MEDLINE=92039036; PubMed=1937019;  
RA Kraetzschmar J., Haendler B., Langer G., Boidol W., Bringmann P.,  
RA Alagon A., Donner P., Schleuning W.D.;  
RT "The plasminogen activator family from the salivary gland of the  
RL vampire bat Desmodus rotundus: cloning and expression.";  
RN Gene 105:229-237(1991).  
RP CHARACTERIZATION.  
RX MEDLINE=93393059; PubMed=1309059;  
RA Schleuning W.-D., Alagon A., Boidol W., Bringmann P., Petri T.,  
RA Kraetzschmar J., Haendler B., Langer G., Baldus B., Witt W.,  
RA Donner P.;  
RT "Plasminogen activators from the saliva of Desmodus rotundus (common  
RL vampire bat): unique fibrin specificity.";  
RN Ann. N.Y. Acad. Sci. 667:395-403(1992).  
RP [3]  
RX X-RAY CRYSTALLOGRAPHY (2.9 ANGSTROMS).  
RC TISSUE=Salivary gland;  
RX MEDLINE=98022741; PubMed=9354616;  
RA Renatus M., Stubbs M.T., Huber R., Bringmann P., Donner P.,  
RA Schleuning W.D., Bode W.;  
RT "Catalytic domain structure of vampire bat plasminogen activator: a  
RL molecular paradigm for proteolysis without activation cleavage.";  
CC Biochemistry 36:13483-13493(1997).  
CC -1- FUNCTION: PROBABLY ESSENTIAL TO SUPPORT THE FEEDING HABITS OF THIS  
CC EXCLUSIVELY HAEMATOPHAGOUS ANIMAL. POTENT THROMBOLYTIC AGENT.  
CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in  
CC plasminogen to form plasmin.  
CC -1- ENZYME REGULATION: ACTIVITY TOWARD PLASMINOGEN IS STIMULATED IN  
CC THE PRESENCE OF FIBRIN I.  
CC -1- SUBUNIT: Monomer.  
CC -1- DOMAIN: THE FIBRONECTIN TYPE-I DOMAIN MEDIATES BINDING TO FIBRIN,  
CC AND THE KRINGLE DOMAIN APPARENTLY MEDIATES FIBRIN-INDUCED  
CC STIMULATION OF ACTIVITY.  
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.  
CC -1- SIMILARITY: Contains 1 EGF-like domain.  
CC -1- SIMILARITY: Contains 1 fibronectin type I domain.  
CC -1- SIMILARITY: Contains 1 kringle domain.  
CC -----  
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; M63987; AAA31591.1; -;  
CC EMBL; M63986; AAA31592.1; -;  
CC PIR; J50597; J50597.  
CC PDB; 1A51; 23-MAR-99.  
CC MEROPS; S01.232; -;  
CC GlycoSuiteDB; P98119; -;  
CC InterPro; IPR001314; Chymotrypsin.  
CC InterPro; IPR006209; EGF like.  
CC InterPro; IPR000083; Fibinctn1.  
CC InterPro; IPR006210; IEGF.  
CC InterPro; IPR000001; Kringle.  
CC InterPro; IPR001254; Ser\_protease\_Try.  
CC Pfam; PF00008; EGF; 1.  
CC Pfam; PF00039; fn1; 1.  
CC Pfam; PF00051; kringle; 1.  
CC Pfam; PF00089; trypsin; 1.  
CC PRINTS; PR00722; CHYMOTRYPSIN.  
CC PRINTS; PR00018; KRINGLE.  
CC ProDom; PD000395; Kringle; 1.  
CC SMART; SM00181; EGF; 1.  
CC SMART; SM00058; FN1; 1.  
CC SMART; SM00130; KR; 1.  
CC SMART; SM00020; Tryp\_Spc; 1.  
CC PROSITE; PS00022; EGF\_1; 1.  
CC PROSITE; PS01186; EGF\_2; 1.









```

DR EMBL; K03226; AAC97138.1; -.
DR EMBL; K02286; AA61252.1; -.
DR EMBL; A21571; CRA01559.1; -.
DR EMBL; A18397; CRA01390.1; -.
DR PIR; A00931; UKHU.
DR PDB; 1KDU; 31-OCT-93.

Query Match      37.7%; Score 777; DB 1; Length 431;
Best Local Similarity 43.3%; Pred. No. 5.6e-60;
Matches 157; Conservative 48; Mismatches 134; Indels 24; Gaps 6;

Qy 29 CYFNGSAYRGTHSLTSGASCLPWNMSILGKVVYTAQNPASOALGLGKHNCRNPDGDA 88
Db 70 CYEGNGHYFKGASTDTMGRCPLPWNSTVLOQTYHAHRSALQGLGKHNCRNPNRR 129
Qy 89 KPWCHLVKNRLTWECYDVPCS-----TCGLRQYSQPFRIKGGFLADIA 134
Db 130 RWCYVQVGLKQVQECWVHCADGKKPSSPPELKCQG-QKTLRPRFKIIGGEFTTIE 188
Qy 135 SHPQAAIAFAKRRSPGRFLCGGILSSCWLSAAHCFQBRFPFPHLTVILGRTYRVP 194
Db 189 NQWFAAIYRRH-RGGSVTVYCGSLSPCWVVISATHCFIDYPKKEDYIVYLGSRSLNSN 247
Qy 195 GEEQKEVEKYVHKFDDDT--YNDIALQLKSDSSRCAQESSVVRTVCLPPADLQL 252
Db 248 TQGMKFEVENLILHKOYSADTLAHNDIALKIRSKGRCAPSRITQITCLPSMYNDP 307
Qy 253 PDWTECELSGKGHEALSPFYSERLKEAHVRLYPSRCTSOHLNRTVTDNMLCAGDTRS 312
Db 308 QFGTSCITGFGKENSITDYLPEQLKMTVVKLISHRECOQPHYGVSEVTTKMLCAD--- 364
Qy 313 GGPOANLHDAQSGGGLVCLNDGRMTLVGIISWGLGCGQKQVPGVYTKVNTYLDWIRD 372
Db 365 --PQWKT-DSQCGSGGLVCLQGRMTLTGIVSGRCALKDKPGVYTVRVSHPFWIRS 421
Qy 373 NNR 375
Db 422 HTK 424

RESULT 11
UROK_PAPCY STANDARD; PRT; 433 AA.
AC P16227;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (uPA)
GN PLAU.
OS Papio cynocephalus (Yellow baboon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecoidea; Papio.
OX NCBI_TaxID=9556;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Thoracic aorta;
RX MEDLINE=90287734; PubMed=2113276;
RA Au Y.P.T., Wang T.W., Clowes A.W.;
RT "Nucleotide and deduced amino acid sequences of baboon urokinase-type
RL plasminogen activator."
RL Nucleic Acids Res. 18:3411-3411(1990).
CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
CC plasminogen to form plasmin.
CC -1- SUBUNIT: FOUND IN HIGH AND LOW MOLECULAR MASS FORMS. EACH CONSISTS
CC OF TWO CHAINS, A AND B. THE HIGH MOLECULAR MASS FORM CONTAINS A
CC LONG CHAIN A. CLEAVAGE OCCURS AFTER RESIDUE 155 IN THE LOW
CC MOLECULAR MASS FORM TO YIELD A SHORT A1 CHAIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1- SIMILARITY: Contains 1 kringle domain.
CC -1- SIMILARITY: Contains 1 EGF-like domain.

```

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

```

CC EMBL; X51935; CAA36200.1; -.
CC PIR; S14687; UKBAY.
CC HSP; P00749; ILWM.
CC MEROPS; S01.231.
CC InterPro; IPR001314; Chymotrypsin.
CC InterPro; IPR006209; EGF like.
CC InterPro; IPR006210; EGF.
CC InterPro; IPR000001; Kringle.
CC InterPro; IPR001254; Ser. protease_Try.
CC Pfam; PF00051; kringle; 1.
CC Pfam; PF00089; trypsin; 1.
CC PRINTS; PR00722; CHYMOTRYPSIN.
CC PRINTS; PR00018; KRINGLE.
CC ProDom; PD000395; Kringle; 1.
CC SMART; SM00181; EGF; 1.
CC SMART; SM00130; KR; 1.
CC SMART; SM00020; Tryp_Spc; 1.
CC PROSITE; PS00022; EGF_1; 1.
CC PROSITE; PS01186; EGF_2; FALSE NEG.
CC PROSITE; PS00021; KRINGLE_1; 1.
CC PROSITE; PS00070; KRINGLE_2; 1.
CC PROSITE; PS00240; TRYPSIN_DOM; 1.
CC PROSITE; PS00134; TRYPSIN_HIS; 1.
CC PROSITE; PS00135; TRYPSIN_SER; 1.
CC Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;
KW Kringle; EGF-like domain; Zymogen; Signal.
FT SIGNAL 1 20
FT CHAIN 21 433
FT CHAIN 21 176
FT CHAIN 155 176
FT CHAIN 178 433
FT DOMAIN 26 62
FT DOMAIN 69 150
FT DOMAIN 151 177
FT DOMAIN 178 433
FT DISULFID 32 50
FT DISULFID 30 38
FT DISULFID 52 61
FT DISULFID 167 298
FT DISULFID 208 224
FT DISULFID 216 287
FT DISULFID 315 384
FT DISULFID 347 363
FT DISULFID 374 402
FT ACT_SITE 223
FT ACT_SITE 274
FT ACT_SITE 378
FT CARBOHYD 324
SQ SEQUENCE 433 AA; 48595 MW; 816D22DFEDDC8792 CRC64;

```

```

Query Match      37.6%; Score 774.5; DB 1; Length 433;
Best Local Similarity 43.5%; Pred. No. 9.2e-60;
Matches 160; Conservative 45; Mismatches 132; Indels 31; Gaps 8;

Qy 29 CYFNGSAYRGTHSLTSGASCLPWNMSILGKVVYTAQNPASOALGLGKHNCRNPDGDA 88
Db 69 CYEGNGHYFKGASTDTMGRCPLPWNSTVLOQTYHAHRSALQGLGKHNCRNPNRR 128
Qy 89 KPWCHLVKNRLTWECYDVPCS-----TCGLRQYSQPFRIKGGFLADIA 134
Db 129 RWCYVQVGLKQVQECWVHCADGKKPSSPPELKCQG-QKTLRPRFKIIGGEFTTIE 187
Qy 135 SHPQAAIAFAKRRSPGRFLCGGILSSCWLSAAHCFQBRFPFPHLTVILGRTYRVP 194
Db 188 NQWFAAIYRRH-RGGSVTVYCGSLSPCWVVISATHCFIDYPKKEDYIVYLGSRSLNSN 246

```

QY 195 GEEQKFEVYVHKFDDDT--YDNDIALQLKSDSSRCAQESSVVRTVCLP-----P 247  
Db 247 TCGENKFEVENILHEDYSADTLAHNDIALKIRSGRCACQPSRTTQTICLPSWYNDP 306  
QY 248 ADLQPLDWTCELSGKGHEALSPFYSERLKEAHLVRLYPSRSTSOHLNLRNTVDNMLCA 307  
Db 307 ND--PPFGTSCEITGKENSSTLYLPEQLKMTVVKLVSQKQCPHYGSEVTTKMLCA 364  
QY 308 GDRSGGQANLHDACGDSGGPLVCLNDGMTLVGLISWGLGCGQKQDVPGYTKVNYL 367  
Db 365 AD-----PQWET--DSCQDGGSGPLVCSIQGHMTLTGIVSWGRGCALKDKPGVYTRVSRFL 418  
QY 368 DWIRDNR 375  
Db 419 PWHSRTR 426

RESULT 12  
UROK RAT STANDARD; PRT; 432 AA.  
AC P29598;  
DT 01-APR-1993 (Rel. 25, Created)  
DT 01-APR-1993 (Rel. 25, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (uPA)  
DE (U-plasminogen activator).  
GN PLAU.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Fischer 344;  
RX MEDLINE=92233409; PubMed=1568219;  
RA Henderson B.R., Tansey W.P., Phillips S.M., Ramshaw I.A.,  
RA Kefford R.F.;  
RT "Transcriptional and posttranscriptional activation of urokinase  
RT plasminogen activator gene expression in metastatic tumor cells.";  
RL Cancer Res. 52:2489-2496(1992).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Kidney;  
RA Rabbani S.A.;  
RL Submitted (APR-1992) to the EMBL/GenBank/DBJ databases.  
CC -!- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in  
CC plasminogen to form plasmin.  
CC -!- SUBUNIT: FOUND IN HIGH AND LOW MOLECULAR MASS FORMS. EACH  
CC CONSISTS OF TWO CHAINS, A AND B. THE HIGH MOLECULAR MASS FORM  
CC CONTAINS A LONG CHAIN A. CLEAVAGE OCCURS AFTER RESIDUE 156 IN THE  
CC LOW MOLECULAR MASS FORM TO YIELD A SHORT A1 CHAIN (BY SIMILARITY).  
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.  
CC -!- SIMILARITY: Contains 1 kringle domain.  
CC -!- SIMILARITY: Contains 1 EGF-like domain.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; X63434; CAA45028.1; -;  
DR EMBL; X65651; CAA46601.1; -;  
DR PIR; S24604; S18932.  
DR HSP; P00749; 1KDU.  
DR MEROPS; S01.231; -;  
DR InterPro; IPR001314; Chymotrypsin.  
DR InterPro; IPR006209; EGF-like.  
DR InterPro; IPR006210; IEGF.  
DR InterPro; IPR000001; Kringle.

DR InterPro; IPR001254; Ser\_protease\_Try.  
DR Pfam; PF00051; kringle; 1.  
DR Pfam; PF00089; trypsin; 1.  
DR PRINTS; PR00722; CHYMOTRYPSIN.  
DR PRINTS; PR00018; KRINGLE.  
DR ProDom; PD000395; Kringle; 1.  
DR SMART; SM00181; EGF; 1.  
DR SMART; SM00130; KR; 1.  
DR SMART; SM00020; Tryp\_SPC; 1.  
DR PROSITE; PS00022; EGF\_1; 1.  
DR PROSITE; PS01186; EGF\_2; FALSE\_NEG.  
DR PROSITE; PS00021; KRINGLE\_1; 1.  
DR PROSITE; PS00070; KRINGLE\_2; 1.  
DR PROSITE; PS00240; TRYPSIN\_DOM; 1.  
DR PROSITE; PS00134; TRYPSIN\_HIS; FALSE\_NEG.  
DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
KW Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;  
KW Kringle; EGF-like domain; Zymogen; Signal.  
FT SIGNAL 1 19  
FT CHAIN 20 432  
FT CHAIN 20 177  
FT CHAIN 156 177  
FT CHAIN 179 432  
FT DOMAIN 27 63  
FT DOMAIN 70 151  
FT DOMAIN 152 178  
FT DOMAIN 179 432  
FT DISULFID 31 39  
FT DISULFID 33 51  
FT DISULFID 53 62  
FT DISULFID 168 300  
FT DISULFID 210 226  
FT DISULFID 218 289  
FT DISULFID 314 383  
FT DISULFID 346 362  
FT DISULFID 373 401  
FT ACT\_SITE 276 276  
FT ACT\_SITE 277 377  
FT ACT\_SITE 16 16  
FT CONFLICT 24 24  
FT CONFLICT 332 332  
SQ SEQUENCE 432 AA; 47957 MW; 4EB1B96C716244C8 CRC64;

Query Match 37.5%; Score 772.5; DB 1; Length 432;  
Best Local Similarity 40.6%; Pred. No. 1.4e-59;  
Matches 147; Conservative 64; Mismatches 128; Indels 23; Gaps 5;

QY 29 CYFNGSAYRGTHSLTESGASCLPWNMSMLIGKVTYTAQNPNSAALGLGKHNCRNPGDA 88  
Db 70 CYHNGSAYRGKANTDTKGRPLAWNSPAVLQQTYNHRSNDSALSLGLGKHNCRNPNQR 129  
QY 89 KPMCHVLNRLTWYCDVPSCS-----TCGLRQYSQPOFRKGLGFADIA 134  
Db 130 RWCYVQIGLKQFVQECWQDCSLSKKPSSTVDQGFQCG-QKALRPRFKVIGGEFTWE 188  
QY 135 SHPQQAIFAHRSPERFLCGGILISSCWLSAAHCFQERFPHHLTVILGRTYRVP 194  
Db 189 NQPMFAAIYLNKGGSPSPFKCGSLISPCWVASATHCFVNQPKKEVYVYLGSKNSY 248  
QY 195 GEEQKFEVYVHKFDDDT--YDNDIALQLKSDSSRCAQESSVVRTVCLPDLQL 252  
Db 249 NPGEMKEVEQLIHEDFSETLAFHNDIALKIRSTGTGCAQPSRTTQTICLPPRGDA 308  
QY 253 PDWTECELSGKGHEALSPFYSERLKEAHLVRLYPSRSTSOHLNLRNTVDNMLCAGTR 312  
Db 309 PFGSDCEITGFGESATDYFYPKDLKMSVVVKIISHEQCKPQHYGSEINVKMLCAAD--- 365  
QY 313 GGPQANLHDACGDSGGPLVCLNDGMTLVGLISWGLGCGQKQDVPGYTKVNYLDIRD 372  
Db 366 --PEWKT--DSCGDSGGGLICNIDGRPTLSGIVSWGSGCAEKKNKPGVYTRVSYPLNIQS 422  
QY 373 NM 374

```

Db      423 HI 424
RESULT 13
UROK_PIG STANDARD; PRT; 442 AA.
AC P04185;
DT 20-MAR-1987 (Rel. 04, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (uPA)
DE (U-plasminogen activator).
GN PLA0.
OS Sus scrofa (pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RX MEDLINE=85087954; PubMed=6096832;
RA Nagamine Y., Pearson D., Altus M.S., Reich E.;
RT "cDNA and gene nucleotide sequence of porcine plasminogen activator.";
RI Nucleic Acids Res. 12:9525-9541(1984).
RN [2]
RP REVISION TO 241.
RA Nagamine Y.;
RL Submitted (DEC-1986) to the PIR data bank.
CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
CC plasminogen to form plasmin.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1- SIMILARITY: Contains 1 kringle domain.
CC -1- SIMILARITY: Contains 1 EGF-like domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
DR EMBL; X01648; CAA25806.1; -.
DR EMBL; X02724; CAA26511.1; -.
DR PIR; A00932; UKPG.
DR HSP; P00749; 1XDU.
DR MEROPS; S01.231; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser protease_Try.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; FALSE NEG.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
DR PROSITE; PS50240; TRYPSIN_HIS; 1.
DR PROSITE; PS00134; TRYPSIN_SER; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;
KW Kringle; EGF-like domain; Zymogen; Signal.
FT SIGNAL 1 20
FT CHAIN 21 442 UROKINASE-TYPE PLASMINOGEN ACTIVATOR.
FT CHAIN 21 188 CHAIN A (BY SIMILARITY).
FT CHAIN 190 442 CHAIN B (BY SIMILARITY).
FT DOMAIN 29 65 EGF-LIKE.

```

---

```

FT DOMAIN 72 153 KRINGLE.
FT DOMAIN 154 189 CONNECTING PEPTIDE.
FT DOMAIN 190 442 SERINE PROTEASE.
FT CARBOHYD 152 152 N-LINKED (GLCNAC. . .).
FT DISULFID 33 41 BY SIMILARITY.
FT DISULFID 35 53 BY SIMILARITY.
FT DISULFID 55 64 BY SIMILARITY.
FT DISULFID 179 310 INTERCHAIN (BY SIMILARITY).
FT DISULFID 220 236 BY SIMILARITY.
FT DISULFID 228 299 BY SIMILARITY.
FT DISULFID 324 393 BY SIMILARITY.
FT DISULFID 356 372 BY SIMILARITY.
FT DISULFID 383 411 BY SIMILARITY.
FT ACT SITE 235 235 CHARGE RELAY SYSTEM.
FT ACT SITE 286 286 CHARGE RELAY SYSTEM.
FT ACT SITE 387 387 CHARGE RELAY SYSTEM.
FT CONFLICT 241 241 Q -> H (IN REF. 1; CAA25806).
FT CONFLICT 242 242 Q -> H (IN REF. 1; CAA26511).
FT CONFLICT 288 288 A -> GS (IN REF. 1; CAA25806).
SQ SEQUENCE 442 AA; 49116 MW; EE32FCEP501321EE CRC64;

Query Match 37.3%; Score 769.5; DB 1; Length 442;
Best Local Similarity 41.2%; Pred. No. 2.6e-59;
Matches 153; Conservative 50; Mismatches 135; Indels 33; Gaps 6;

Qy 29 CYFGNGSAYRGTHSLTESGASCLPWNMSILIGKYVTAQNPSAQALGLGKHNYCRNPDGA 88
Db 72 CFEGNGHSYRGKANTNGRPLPWN SATVLLNTYHAHRPDALQLGLGKHNYCRNPDQR 131
Qy 89 KPCHVLKNRLTWECYDVSCS-----TCGLRQYQOPQPRI 125
Db 132 RPMCYYQVGLKQLVQECWPNCSGGESHRRPAYDKNPFSTPEKVFQCG-QKALRRPFI 190
Qy 126 KGLFLADIASHPQAAIFAKHRRSPGRFLCGGLISSCWLSAAHCFQERFPHLTVI 185
Db 191 VGGKSTTIENQPAALYRRH-RGGSVTVCGGSLSPCWVSVATHCFINYYQKEDYIVY 249
Qy 186 LGRTYRVVPEEEQKFEVKYIVHKEFDDDT--YDNDIALQLKSDSSRCAQESSVVRVY 243
Db 250 LGRQTLSSHTGEMKFEVKILHEDYSADSLAHHNDIALKIRTDKGCAQPSRSIQTI 309
Qy 244 CLPPADQLPDWTECELSGYGKHEALSPFFYSERLKEAHRVLYPSRCTSHLLNRTVTDN 303
Db 310 CLPPVNGDAHFGASCEIVGFGKEDPSDYLPQLKMTVVKLVSHRECCQPHYYGSEVTTK 369
Qy 304 MLCAGDTRSGGPOANLHDACQSGGGLVCLNDGRMTLVGIISWGLCGGKQKDVGVVTKV 363
Db 370 MLCAAD-----POWKT-DSQGGSGGGLVCSGTQRLTLTGIVSWGRECAKMDKPGVYTRV 423
Qy 364 TNYLDWIRDNM 374
Db 424 SRFLTWLHTHV 434

RESULT 14
UROK_CHICK
ID UROK_CHICK STANDARD; PRT; 434 AA.
AC P15120;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (uPA)
DE (U-plasminogen activator).
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90110185; PubMed=2295632;
RA Leslie N.D., Kessler C.A., Bell S.M., Degen J.L.;
RT "The chicken urokinase-type plasminogen activator gene.";

```

Query Match 37.0%; Score 762; DB 1; Length 434;  
Best Local Similarity 45.4%; Pred.No.1.le-58;  
Matches 163; Conservative 56; Mismatches 116; Indels 24; Gaps 12;

QY 26 NSDCYFGNSAVRGTHTSLTSGASCIPWN--SMILIGKVYTAQNPSAQLGLGKHNRCRN 83  
| | | | | | | | | | | | | | | | | | | | | | | | | | | |  
Db 76 NSICYSGNGEDVRG----NAEDPGCLYWDPHSVIRWGD-YHADLKVALGLGKHNRCRN 130

Search completed: August 8, 2003, 17:14:13  
Job time : 17 secs

```
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; FALSE_NEG.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW plasminogen activation; Hydrolase; Serine protease; Glycoprotein;
KW Kringle; EGF-like domain; Signal; Zymogen.
FT SIGNAL 1 20
FT CHAIN 21 433 UROKINASE-TYPE PLASMINOGEN ACTIVATOR.
FT CHAIN 21 179 CHAIN A (BY SIMILARITY).
FT CHAIN 181 433 CHAIN B (BY SIMILARITY).
FT DOMAIN 29 65 EGF-LIKE.
FT DOMAIN 72 153 KRINGLE.
FT DOMAIN 154 180 CONNECTING PEPTIDE.
FT DOMAIN 181 433 SERINE PROTEASE.
FT DISULFID 33 41 BY SIMILARITY.
FT DISULFID 35 53 BY SIMILARITY.
FT DISULFID 55 64 BY SIMILARITY.
FT DISULFID 170 301 INTERCHAIN (BY SIMILARITY).
FT DISULFID 211 227 BY SIMILARITY.
FT DISULFID 219 290 BY SIMILARITY.
FT DISULFID 315 384 BY SIMILARITY.
FT DISULFID 347 363 BY SIMILARITY.
FT DISULFID 374 402 BY SIMILARITY.
FT ACT_SITE 226 226 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 277 277 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 378 378 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT CONFLICT 189 189 A -> T (IN REF. 2).
SQ SEQUENCE 433 AA; 48730 MW; 4DE1B8D4DA47027A CRC64;

Query Match 36.1%; Score 745; DB 1; Length 433;
Best Local Similarity 41.3%; Pred. No. 3.4e-57;
Matches 150; Conservative 54; Mismatches 135; Indels 24; Gaps 6;

Qy 29 CYFGNGSAYRCHSLTSGASCLPWNMSILIGKYTAQNPSAQALGLGKHNKCNPDGDA 88
Db |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
72 CYQNGHSYRGKANDLSGRPLCDAWSDPTVLLKMYHAHRSDAIQLGLGKHNKCNPNQR 131
Qy 89 KPWCHVLKNRLTWECYDVPCST-----CGLRQYSQPFRIKGLFADIA 134
Db |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
132 RWCYVQIGLQFVQFCWQDCVSGKSPSPREKEFCQG-QKALRPRFKIVGGQVTNAE 190
Qy 135 SHPQAAIFAKHRSPGFRFLCGGILISSCWILSAACFQRPFPHLTVILGRTRYVP 194
Db :||||:||||:||||:||||:||||:||||:||||:||||:||||:
191 NQPFAAIYRRH-RGGSTYLCGSLISPCWWSATHCFIDHPKKENIVYVVGQRLNSD 249
Qy 195 GEEQKEVEKYIVHKSFDDDT--YDNDIALQLKSDSSRCAOESSVVRTVCLPPADLQL 252
Db |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
250 TRGEMQFEVEKLIHEDYSAESLAHNDIALLKIRTSRGCAQPSRSIQITCLPEHEDA 309
Qy 253 PDWTECLSGYGKHEALSPFYSERLKEAHVRLYPSRCTSQHLNRTVTNNMLCAGDTRS 312
Db |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
310 HSRTRCEITGKPNPSDYRSDCLKMTFVSLVSHEVCQPHYYGAETDKMLCAAD--- 366
Qy 313 GGPQANLHDAQCGSGGRLVCLNDRMTLVGIIISWGLGCCQKQVPGVYTKVNTYLDWIRD 372
Db |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
367 --PQWET-DSQQGDSGGPLVCTIQGLRTLGTGIVSGRDCAMKYKPGVYTVRSKELPWINT 423
Qy 373 NMR 375
Db :|
424 HTR 426
```

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 8, 2003, 17:11:32 ; Search time 39 Seconds  
(without alignments)  
2494.509 Million cell updates/sec

Title: US-09-987-455-8

Perfect score: 2061

Sequence: 1 MKTATIAIAVALAGFATVAQ.....GVYTKVTNYLDWIRDNNRPG 377

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL\_23:\*\*

- 1: sp\_archaea:\*\*
- 2: sp\_bacteria:\*\*
- 3: sp\_fungi:\*\*
- 4: sp\_human:\*\*
- 5: sp\_invertebrate:\*\*
- 6: sp\_mammal:\*\*
- 7: sp\_mhc:\*\*
- 8: sp\_organelle:\*\*
- 9: sp\_phase:\*\*
- 10: sp\_plant:\*\*
- 11: sp\_rodent:\*\*
- 12: sp\_virus:\*\*
- 13: sp\_vertebrate:\*\*
- 14: sp\_unclassified:\*\*
- 15: sp\_rvirus:\*\*
- 16: sp\_bacteriap:\*\*
- 17: sp\_archaeap:\*\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1961	95.1	516	4 Q8BU99	Q8BU99 homo sapien
2	1942	94.2	395	4 Q8BZW1	Q8BZW1 homo sapien
3	1669	81.0	562	6 Q8SQ23	Q8SQ23 sus scrofa
4	1601	77.7	559	11 Q91VP2	Q91VP2 mus musculus
5	1536	74.5	564	6 Q8MKB1	Q8MKB1 oryctolagus
6	769	37.3	433	6 Q8MIL0	Q8MIL0 oryctolagus
7	767	37.2	433	6 Q8MHV7	Q8MHV7 oryctolagus
8	763	37.0	202	13 Q90675	Q90675 gallus gall
9	742.5	36.0	653	11 Q8VCS4	Q8VCS4 mus musculus
10	709.5	34.4	560	4 Q14520	Q14520 homo sapien
11	698.5	33.9	517	11 Q8K0D2	Q8K0D2 mus musculus
12	667.5	32.4	615	4 Q8I2Z5	Q8I2Z5 homo sapien
13	634.5	30.8	597	11 Q35727	Q35727 mus musculus
14	630	30.6	616	6 Q97507	Q97507 sus scrofa
15	607	29.5	812	11 Q91WJ5	Q91WJ5 mus musculus
16	593.5	28.8	812	11 Q9ROW3	Q9ROW3 rattus norv

17	580	28.1	806	6	O18783
18	560.5	27.2	810	4	Q15146
19	555.5	27.0	334	6	O46507
20	549	26.6	454	6	O46506
21	525.5	25.5	429	13	Q8AVB0
22	487.5	23.7	300	4	Q96EF3
23	463.5	22.5	385	5	Q25101
24	453	22.0	855	11	Q9JJ17
25	452	21.9	761	11	Q99JC8
26	449	21.8	422	4	Q8WVC1
27	443.5	21.5	90	4	Q8NG20
28	443.5	21.5	802	4	Q8IUE2
29	443.5	21.5	811	4	Q8IU80
30	442.5	21.5	868	5	Q9YIV3
31	433.5	21.0	608	13	Q9PTW7
32	427.5	20.7	799	11	Q9DBI0
33	425.5	20.6	214	6	Q9XT70
34	423.5	20.5	607	13	Q91001
35	422.5	20.5	267	5	Q9BK47
36	422	20.5	371	11	Q8CJ16
37	416.5	20.2	445	11	Q8CJ17
38	416.5	20.2	471	11	Q8CFE0
39	415.5	20.2	505	5	Q96SV4
40	409.5	19.9	537	4	Q9BYE1
41	409.5	19.9	581	4	Q9BYE2
42	408	19.8	455	11	Q8CDR0
43	407.5	19.8	394	5	P91817
44	405	19.7	613	13	Q03711
45	401	19.5	453	11	Q8VDE0

## ALIGNMENTS

### RESULT 1

Q8BU99	PRELIMINARY;	PRT;	516 AA.
ID	Q8BU99		
AC	Q8BU99;		
DT	01-JUN-2001 (Tremblrel. 17, Created)		
DT	01-JUN-2001 (Tremblrel. 17, Last sequence update)		
DT	01-MAR-2003 (Tremblrel. 23, Last annotation update)		
DE	Similar to plasminogen activator, tissue.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
OX	NCBI_TaxID=9606;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=Skin;		
RA	Strausberg R.;		
RL	Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.		
CC	-1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.		
CC	-1- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.		
DR	EMBL; BC02795; AA02795.1; -		
DR	HSSP; P00750; 1A5H.		
DR	InterPro; IPR001314; Chymotrypsin.		
DR	InterPro; IPR006209; EGF like.		
DR	InterPro; IPR006210; IEGF.		
DR	InterPro; IPR000001; Kringle.		
DR	InterPro; IPR001254; Ser_protease_Try.		
DR	Pfam; PF00008; EGF; 1.		
DR	Pfam; PF00051; kringle; 2.		
DR	Pfam; PF00089; tryptain; 1.		
DR	PRINTS; PR00722; CHYMOTRYPSIN.		
DR	PRINTS; PR00018; KRINGLE.		
DR	ProDom; PD000395; Kringle; 2.		
DR	SMART; SM00181; EGF; 1.		
DR	SMART; SM00130; KR; 2.		
DR	SMART; SM00020; Tryp_SPC; 1.		
DR	PROSITE; PS00022; EGF_1; 1.		
DR	PROSITE; PS01186; EGF_2; 1.		
DR	PROSITE; PS00021; KRINGLE_1; 2.		
DR	PROSITE; PS00070; KRINGLE_2; 2.		

[illegible]

```

RESULT 3
Q8SQ23 PRELIMINARY; PRT; 562 AA.
AC Q8SQ23;
DT 01-JUN-2002 (T-EMBLrel. 21, Created)
DT 01-JUN-2002 (T-EMBLrel. 21, Last sequence update)
DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
DT T-plasminogen activator.
OS Sus scrofa [pig].
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
NCBI_TaxId=9623;
RN [1]
SEQUENCE FROM N.A.
RT TISSUE=Enamel organ.
RA Ding Y., Xue J., Bartlett J.D.;
RT "T-plasminogen activator in tooth tissues";
RL Submitted (MAR-2001) to the EMBL/Genbank/DBJ databases.
CC 1- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.
DR EMBL; AF364605; AAM00297.1; -.
DR HSP; P00761; IAN1.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR000083; Fib_rnctnl.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00008; EGF; 1.
DR Pfam; PF00039; fn1; 1.
DR Pfam; PF00051; kringle; 2.
DR Pfam; PF00089; trypsin; 1.
DR ProDom; PD000395; Kringle; 2.
DR SMART; SM00181; EGF; 1.
DR SMART; SM00058; FN1; 1.
DR SMART; SM00130; KR; 2.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; 1.

```

DR PROSITE; PS01253; FIBRONECTIN\_1; 1.  
 DR PROSITE; PS00021; KRINGLE\_1; 2.  
 DR PROSITE; PS50070; KRINGLE\_2; 2.  
 DR PROSITE; PS50240; TRYPSIN\_DOM; 1.  
 DR PROSITE; PS00134; TRYPSIN\_HIS; 1.  
 DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
 KW EGF-like domain; Glycoprotein; Hydrolase; Kringle; Protease;  
 KW Serine protease.  
 SQ SEQUENCE 562 AA; 63668 MW; F9E6B4C77CB101E8 CRC64;

Query Match 81.0%; Score 1669; DB 6; Length 562;  
 Best Local Similarity 82.0%; Pred. No. 8.9e-158;  
 Matches 292; Conservative 27; Mismatches 37; Indels 0; Gaps 0;

QY 21 AASEGSDCYFGNGSAYRGTHSLTESGASCLPWNMSMLIGKVYTAQNPQAALGLGKHNY 80  
 DB 207 ACTKEKECYGTGKGLDYGRTRSLTMSGAFCLPWNLSVLGMKIYTAWNSNAQTGLGKHNY 266

QY 81 CRNPDGDAKPCHVLKNRRLTWECYCDVPSCTGRLQYSDQFRIKGLFADIASHPWA 140  
 DB 267 CRNPDGDPQPCWCHVKDKHLTWECYCDVPSCTGRLQYSDQFRIKGLFADIASHPWA 326

QY 141 AIFAKHRRSPGERFLCGGILISSCWLSAAHCFORPPPHLTLVILGRTYRVVPGEEQK 200  
 DB 327 AIFVKNRRSPGERFLCGGILISSCWLSAAHCFORPPPHVVRVGLGRTYRVVPGEEQK 386

QY 201 FEVEKYIVHKEFDDDDTYNDIALQLKSDSRCAQESSVVRVTCVLPADLQLPDWTCE 260  
 DB 387 FEVEKYIVHKEFDDDDTYNDIALQLKSDSLTCAQESDAVTVCLPEANLQLPDWTCE 446

QY 261 SGYKGHEALSPFYSERLKEAHRVLPSSRCTSOHLNRTVTNNMLCAGDTRSGGQANLH 320  
 DB 447 SGYKGHEALSPFYSERLKEAHRVLPSSRCTSKHLFNKTIANNMLCAGDTRSGGQANLH 506

QY 321 DACQSDGGLVCLNDGRMTLVGIISWGLGCGQKDPVGVYTKVTNYLWDIRDNR 376  
 DB 507 DACQSDGGLVCLNMGNHTLVGIISWGLGCGQKDPVGVYTKVTNYLWDIRDNR 562

RESULT 4  
 Q91VP2  
 ID Q91VP2 PRELIMINARY; PRT; 559 AA.  
 AC Q91VP2  
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE Similar to plasminogen activator, tissue.  
 GN PLAT.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Breast tumor;  
 RA Strausberg R.;  
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.  
 CC -1- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.  
 DR EMBL; BC011256; AAL11256.1; -;  
 DR HSSP; P00761; IAN1.  
 DR MGD; MGI:97610; Plat.  
 DR InterPro; IPR001314; Chymotrypsin.  
 DR InterPro; IPR006209; EGF like  
 DR InterPro; IPR000083; Fibrinctnl.  
 DR InterPro; IPR006210; IEGF.  
 DR InterPro; IPR000001; Kringle.  
 DR InterPro; IPR001254; Ser\_protease\_Try.  
 DR Pfam; PF00008; EGF; 1.  
 DR Pfam; PF00039; fn1; 1.  
 DR Pfam; PF00051; kringle; 2.  
 DR Pfam; PF00089; trypsin; 1.  
 DR PRINTS; PR00722; CHYMOTRYPSIN.

DR PRINTS; PR00018; KRINGLE.  
 DR ProDom; PD000395; Kringle; 2.  
 DR SMART; SM00181; EGF; 1.  
 DR SMART; SM00058; FN1; 1.  
 DR SMART; SM00130; KR; 2.  
 DR SMART; SM00020; Tryp\_SPC; 1.  
 DR PROSITE; PS00022; EGF\_1; 1.  
 DR PROSITE; PS01186; EGF\_2; 1.  
 DR PROSITE; PS01253; FIBRONECTIN\_1; 1.  
 DR PROSITE; PS00021; KRINGLE\_1; 2.  
 DR PROSITE; PS50070; KRINGLE\_2; 2.  
 DR PROSITE; PS50240; TRYPSIN\_DOM; 1.  
 DR PROSITE; PS00134; TRYPSIN\_HIS; 1.  
 DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
 KW EGF-like domain; Glycoprotein; Hydrolase; Kringle; Protease;  
 KW Serine protease.  
 SQ SEQUENCE 559 AA; 63122 MW; 8CCBE2BDB94514D9 CRC64;

Query Match 77.7%; Score 1601; DB 11; Length 559;  
 Best Local Similarity 79.5%; Pred. No. 5.5e-151;  
 Matches 283; Conservative 34; Mismatches 37; Indels 2; Gaps 2;

QY 21 AASEGNS-DCVFGNGSAYRGTHSLTESGASCLPWNMSMLIGKVYTAQNPQAALGLGKH 79  
 DB 204 ACPGKSEDCYVGKGYRTGTHSLTTSQASCLPWNLSVLGMKSYTAWRTSQALGLGRH 263

QY 80 YCRNPDGDAKPCHVLKNRRLTWECYCDVPSCTGRLQYSDQFRIKGLFADIASHPW 139  
 DB 264 YCRNPDGDAKPCHVLKNRRLTWECYCDVPSCTGRLQYSDQFRIKGLFADIASHPW 323

QY 140 AIFAKHRRSPGERFLCGGILISSCWLSAAHCFORPPPHLTLVILGRTYRVVPGEEQ 199  
 DB 324 AIFVKNRRSPGERFLCGGILISSCWLSAAHCFORPPPHLTLVILGRTYRVVPGEEQ 383

QY 200 KFEVEKYIVHKEFDDDDTYNDIALQLKSDSRCAQESSVVRVTCVLPADLQLPDWTCE 259  
 DB 384 TFEIEKYIVHKEFDDDDTYNDIALQLKSDSRCAQESSVVRVTCVLPADLQLPDWTCE 443

QY 260 LSGYKGHEALSPFYSERLKEAHRVLPSSRCTSOHLNRTVTNNMLCAGDTRSGGQANL 319  
 DB 444 LSGYKGHEALSPFYSERLKEAHRVLPSSRCTSOHLNRTVTNNMLCAGDTRSGGQANL 502

QY 320 HDACQSDGGLVCLNDGRMTLVGIISWGLGCGQKDPVGVYTKVTNYLWDIRDNR 375  
 DB 503 HDACQSDGGLVCLNMGNHTLVGIISWGLGCGQKDPVGVYTKVTNYLWDIRDNR 558

RESULT 5  
 Q8MKB1  
 ID Q8MKB1 PRELIMINARY; PRT; 564 AA.  
 AC Q8MKB1  
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE Tissue-type plasminogen activator.  
 OS Oryctolagus cuniculus (Rabbit).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
 OX NCBI\_TaxID=9986;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Sugiki M., Yoshida E., Anai K., Maruyama M.;  
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.  
 CC -1- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.  
 DR EMBL; AY029518; AAK40240.1; -;  
 DR InterPro; IPR001314; Chymotrypsin.  
 DR InterPro; IPR006209; EGF like.  
 DR InterPro; IPR000083; Fibrinctnl.  
 DR InterPro; IPR006210; IEGF.  
 DR InterPro; IPR000001; Kringle.  
 DR InterPro; IPR001254; Ser\_protease\_Try.  
 DR Pfam; PF00008; EGF; 1.



```

[1]
RP SEQUENCE FROM N.A.
RA Sugiki M., Yoshida E., Anai K., Maruyama M.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
[2]
RP SEQUENCE FROM N.A.
RA Yano W., Watanabe M.;
RT "Oryctolagus cuniculus urokinase-type plasminogen activator, mRNA,
complete cds."
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; AY029517; AAK40239.1; -.
DR EMBL; AB087224; BAC02685.1; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser.protease_Try.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS50070; KRINGLE_2; 1.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Glycoprotein; Hydrolase; Kinase; Kringle; Protease; Serine protease.
SQ SEQUENCE 433 AA; 48444 MW; 6DD35A371010A6EE CRC64;

Query Match 37.2%; Score 767; DB 6; Length 433;
Best Local Similarity 43.1%; Pred. No. 7.8e-68;
Matches 157; Conservative 52; Mismatches 127; Indels 28; Gaps 8;

Qy 29 CYFGNGSAYRGTHSLTESGASCLPWNMILIGKVYTAQNPSAQAALGLGKHNCRNPGDA 88
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 89 KPWCHLVKNRLTWECYDVPSCST-----CGLRQYSQPFRIKGGFLADIA 134
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 132 RPWCYVQVGLQLIQECKVHDCSSGKKPALPPGKLEFCQG-QKALRPFRKIIGFEFTIE 190
Qy 135 SHPQQAIFAKHRSRPGERFLCGGILISSCWILSAHCFQERFPFPHLTVILGRT--YRV 192
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 191 NQFWAALYRRH-RGGSVTYVCGSLISPCWVVSATHCFINHQKEDYIVYLGSRSLNSM 249
Qy 193 VPGEEQKFEYKIVHKEFDDET--YDNDIALQLKSDSRCAQESSVVRVTCVLPADL 250
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 250 TPG--EMKFEVQLIHGYRADTLAHNDIALKILSNNGCAQPSRSIGTICLPWNA 307
Qy 251 QLPWTEBCELSGYGHEALSPFYSERLKEAHLRYPSSRCTSQHLNRTVTDNMLCAGDT 310
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 308 DPNFGTSCETFGKENSTDYLPQLKWTVVKLVSYQEQCPHYGSEVTTKMLCAAD- 366
Qy 311 RSGGQANLHACCGDSGGLVCLNDGMTLVGLISWGLCGQKQDVGVYTKVTVNLDWI 370
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 367 ----PQWET-DSCQDGGGLVCSVQGRMTLTGIVSWGRGALKKPKGVYTRVSRFLPW 421
Qy 371 RDNM 374
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 422 RSHI 425

RESULT 8
ID Q90675 PRELIMINARY; PRT; 202 AA.
AC Q90675;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DR 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)

DT DE 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
GN Tissue-type plasminogen activator (Fragment).
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=White Leghorn;
RX MEDLINE=97199025; Pubmed=9047000;
RA Johnson A.L., Bridgham J.T., Anthony R.V.;
RT "Expression of avian urokinase and tissue-type plasminogen activator
messenger ribonucleic acid during follicle development and atresia.";
RL Biol. Reprod. 56:581-588(1997).
CC -1- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.
DR EMBL; U31988; AAA74955.1; -.
DR HSP; P00750; IRTF.
DR MEROPS; S01.232; -.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser.protease_Try.
DR Pfam; PF00051; kringle; 2.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 2.
DR SMART; SM00130; KR; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS50070; KRINGLE_2; 1.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Glycoprotein; Hydrolase; Kinase; Kringle; Protease; Serine protease.
FT NON_TER 1
FT NON_TER 202
SQ SEQUENCE 202 AA; 23148 MW; 049DB42941D83AFC CRC64;

Query Match 37.0%; Score 763; DB 13; Length 202;
Best Local Similarity 77.4%; Pred. No. 7.1e-68;
Matches 130; Conservative 16; Mismatches 22; Indels 0; Gaps 0;

Qy 23 SEGSDCYFGNGSAYRGTHSLTESGASCLPWNMILIGKVYTAQNPSAQAALGLGKHNCR 82
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 83 NPQGDAPWCHLVKNRLTWECYDVPSCSTCGLRQYSQPFRIKGGFLADIAHPWQAAI 142
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 95 NPQGDAPWCHLVKNRLTWECYDVPSCSTCGLRQYSQPFRIKGGFLADIAHPWQAAI 154
Qy 143 FAXHRSRPGERFLCGGILISSCWILSAHCFQERFPFPHLTVILGRTY 190
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 155 FVKNRAPGQRFLLCGGILISSCWILSAHCFQERFPFPHLTVILGRTY 202

RESULT 9
ID Q8VCS4 PRELIMINARY; PRT; 653 AA.
AC Q8VCS4;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical 70.6 kDa protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Strausberg R.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; BC019376; AAH19376.1; -.

```



207	DG	DGDKWCFVKVNSEKVKWEYCDVTVCPDPTNPVSVLLEPVMELPGFSSCGKTEVAEH	266			
122	QF	-RIKGLPADIASHPWQAIAFAK--HRRSPGERFLCGGILISSCWILSAAHCFQBRFP	178			
267	AV	KRIYGGFKSTAGKHPQWVSLQTSPLTTSMQGHFCGALHPCWVLTAAHCTD--IN	324			
179	PH	LTVLGRYRVVPGEBEKEVEKYIVHKBF--DITYNDIALLOLKSDSSRCAQE	236			
325	TE	LKVVLDQDLKTKTESHEQTFERVEKILKYSQYNERDEIPHNDIALLKLPVGGHCALE	384			
237	SS	VTVTPPADLQLPDWTCELSGVGKHEALSFPYSERLKEAHVRLYPSSRCTSOHL	296			
385	SR	VKTVCCL-PSD-PPFSGTECHISGMVTVETGEG--SROLLDAKVLIANPLCNSQLY	440			
297	NR	TVDNMLCAGDTRSGGPOANLHDAQQDSGGPLVCLNDGRMTLVGIIISWGLCGCGKDV	356			
441	DT	IDSIMICAGNLQKPG-----SDTCQDSGGPLCEKDKGTYVYVGVSWGQECGKK--	493			
357	PG	VYTKVTVNLYDWRDNM	374			
494	PG	VTVTKVTKFLNKKITM	511			
RESULT 12						
Q81Z25 PRELIMINARY; PRT; 615 AA.						
ID	Q81Z25	AC	Q81Z25;			
DT	01-MAR-2003	(T-EMBLrel. 23, Created)				
DT	01-MAR-2003	(T-EMBLrel. 23, Last sequence update)				
DT	01-MAR-2003	(T-EMBLrel. 23, Last annotation update)				
OS	Coagulation factor XII-Mie.					
DS	Homo sapiens (Human).					
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;					
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.					
OX	NCBI_TaxID=9606;					
RP	[1]					
RN	SEQUENCE FROM N. A.					
RA	Wada H., Nishioaka J., Nakatani K., Kasai Y., Abe Y., Nobori T.;					
RT	"Molecular characterization of coagulation factor XII-Mie.;"					
RL	Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.					
DR	EMBL;	AB095845; BAC23095.1; --				
SQ	SEQUENCE 615 AA; 67735 MW; 030508870AC07EDB CRC64;					
Query Match 32.4%; Score 667.5; DB 4; Length 615;						
Best Local Similarity 36.2%; Pred. No. 1.1e-57;						
Matches 150; Conservative 48; Mismatches 133; Indels 83; Gaps 7;						
QY	27	SDCYFNGSAYRGTHTSLTESGASCLPWNMILGKVVTAQNPQAQALGLGKHNYCRNPDG	86			
DB	215	ASCYDGRGLSYRGLARTLTLSGAPCQPWASEATYRNVTAEQ---ARNWGLGCHAFCRNPDN	271			
QY	87	DAKPNCHVLKNRLTWECYDCVPSCST-----	112			
DB	272	DIRPWCFLNRDLRSWEYCDLAQCQTPTAAPTTPVSLHVLMPAQAAPPKQPPTRT	331			
QY	113	-----CG--LRQYSQPQPRIKGLFADIASHPWQAIF	143			
DB	332	PPQSQTGALPAKREQPPSLTRNGPLSCGQRLAKSLSSMTRVVGLVLRGAHPYIALLY	391			
QY	144	AKHRRSPGERFLCGGILISSCWILSAAHCFQBRPPPHLTVILGRYRVVPGEBEKEVEKY	203			
DB	392	WGHS-----FCAGSLIAPCWVLTAAHCLQDRPAPEDLTVVLGQERRNHSCEPQTLAV	444			
QY	204	EKIVHKEPDDTYNDIALLOLKSDS-SRCAQESSVVRVTVCLPPADLQLPDWTCELSG	262			
DB	445	RSYRLHEAFSPVQHDALLRLQEDADGSCALLSPYQVQPLPSGNAARESETTLCOVAG	504			
QY	263	YGRKEALSPPYSERLKEAHVRLYPSSRCTSOHLNRTVTDNMLCAGDTRSGGPOANLHDA	322			
DB	505	CGHQFEGAEEYASFLQEAQVFPFLSLERCSAPDVHGSILPGMLCAGFLEGG-----TDA	558			
QY	323	COQDSGGPLVCLN---DGRMTLVGIIISWGLCGCGKQVPGYVTKVTVNLYDWRDNM	373			

```
Db 559 CQDGGGLVCEBQAAERLLTQIISWGGCGDRNKPVGVTVDVAYLAWIREH 612

RESULT 13
O35727 PRELIMINARY; PRT; 597 AA.
AC O35727;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Factor XII.
GN F12.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Schloesser M., Schwager S., Engel W.;
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
DR EMBL; X99571; CAA67891.1; -.
DR HSP; P00760; IAO7.
DR MEROPS; S01.211; -.
DR MGD; MGI:1891012; F12.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR000083; Fibnctnl.
DR InterPro; IPR000562; FN Type II.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00008; EGF; 2.
DR Pfam; PF00039; fn1; 1.
DR Pfam; PF00040; fn2; 1.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00013; FNTYPEII.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000995; FN Type II; 1.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00181; EGF; 2.
DR SMART; SM00058; FN1; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_SPC; 1.
DR SMART; SM00059; FN2; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS00022; EGF_1; 2.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01253; FIBRONECTIN_1; 1.
DR PROSITE; PS00023; FIBRONECTIN_2; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS0240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW EGF-like domain; Glycoprotein; Hydrolase; Kringle; Protease;
SQ Serine protease.
SQ SEQUENCE 597 AA; 65638 MW; F3AC07C37D0C0FBA CRC64;

Query Match 30.8%; Score 634.5; DB 11; Length 597;
Best Local Similarity 36.4%; Pred. No. 2.1e-54;
Matches 144; Conservative 48; Mismatches 137; Indels 67; Gaps 9;

Qy 29 CYFGNGSYRGTHSLTESGASCLPWNMILIGKVY-TAQNPSAALGLGKHYCRNPNGD 87
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 217 CYEGGLSYRGAGTQSGACQRTW----VEATYRNTEKQALSWGGLGHAFCEPNDD 272
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 88 AKPWCHLVNRLRWTEYCDVSCST-----
|||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 273 TRPWCFFVMSGDRLSWDYCGLEQCQTPTFAPLVWPESQSPSPSLSHAPNDSTDHQT 332
|||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 113 -----CG--LROYSPQPFRIKGGIFADIASHPQAOAIFAKHRRSPGERFLCGGILIS 162
|||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 333 LSKNTNMGCCQRFKGLSSFMRVVGGVLVALPGSHPIAALYW-----GNF-CAGSLIA 385
|||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 163 SCWLSAAHCQBFRRFPFPHLTVILGRTRYVVPGESEKFEVEKYIVHKFDDDDTDYNDIA 222
|||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 386 PCWVLTAAHCLQNRPAPEELTVVLGQDRHNQSCWQCQLAVRSYRLHEGFSSTIYQHDLA 445
|||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 223 LLQI-KSDSRCAQESSVVTCLPPADLQLPDWTCELSGYKHEALSPFYSRLKEAH 281
|||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 446 LLRLQESKTNSCAILSPHVQVCLPSGAAPPSETVLCEVAGWGHLGAEYEYTFLOEAQ 505
|||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 282 VRLYPSSRCTSQHLNRTVTDMLCAGDTRSGGPQANLHDACQDGGPLVC---LNDGR 338
|||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 506 VPFIALDRCSNSNVHGDAILPGMLCAGFLEGG-----TDACQDGGPLVCBEGTAHQ 559
|||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 339 MTLVGIISWGLGCGQKQDVPGVYTKVTNYLDWIRDNM 374
|||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 560 LTLRGVLSWGGCGDRNKPVGVTVDVAYLAWIQHI 595
|||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 14
O97507 PRELIMINARY; PRT; 616 AA.
AC O97507;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE FXII.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Takahashi T., Kihara T.;
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
DR EMBL; AB024246; BAA37148.1; -.
DR HSP; P00763; IDPO.
DR MEROPS; S01.211; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000742; EGF_2.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR000083; Fibnctnl.
DR InterPro; IPR000562; FN Type II.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00008; EGF; 2.
DR Pfam; PF00039; fn1; 1.
DR Pfam; PF00040; fn2; 1.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00013; FNTYPEII.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000995; FN Type II; 1.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00181; EGF; 2.
DR SMART; SM00058; FN1; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS00022; EGF_1; 2.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01253; FIBRONECTIN_1; 1.
DR PROSITE; PS00023; FIBRONECTIN_2; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS0240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW EGF-like domain; Glycoprotein; Hydrolase; Kringle; Protease;
SQ Serine protease.
SQ SEQUENCE 616 AA; 65638 MW; F3AC07C37D0C0FBA CRC64;
```



[illegible]

Search completed: August 8, 2003, 17:15:00  
Job time : 41 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model  
Run on: August 18, 2003, 19:53:48 ; Search time 277.124 Seconds  
(without alignments)  
9743.047 Million cell updates/sec

Title: US-09-987-455-3  
Perfect score: 66  
Sequence: 1 atgaaaagacagtcagc.....ctacgtggccagcgcc 66

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl:

- 1: gb\_bat:
- 2: gb\_htg:
- 3: gb\_in:
- 4: gb\_om:
- 5: gb\_ov:
- 6: gb\_pat:
- 7: gb\_ph:
- 8: gb\_pi:
- 9: gb\_pr:
- 10: gb\_ro:
- 11: gb\_sts:
- 12: gb\_sy:
- 13: gb\_un:
- 14: gb\_vl:
- 15: em\_ba:
- 16: em\_fun:
- 17: em\_hum:
- 18: em\_in:
- 19: em\_mu:
- 20: em\_om:
- 21: em\_or:
- 22: em\_ov:
- 23: em\_pat:
- 24: em\_ph:
- 25: em\_pl:
- 26: em\_ro:
- 27: em\_sts:
- 28: em\_un:
- 29: em\_vl:
- 30: em\_htg\_hum:
- 31: em\_htg\_inv:
- 32: em\_htg\_other:
- 33: em\_htg\_mus:
- 34: em\_htg\_pln:
- 35: em\_htg\_rod:
- 36: em\_htg\_mam:
- 37: em\_htg\_vrt:
- 38: em\_sy:
- 39: em\_htgo\_hum:
- 40: em\_htgo\_mus:
- 41: em\_htgo\_other:

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	66	100.0	66	6	AX430829	Sequence
2	66	100.0	66	6	AX431269	Sequence
3	66	100.0	66	6	AX431272	Sequence
4	66	100.0	705	12	HSIGKLC28	Synthetic c
5	66	100.0	705	12	HSIGKLC5	Synthetic c
6	66	100.0	711	12	HSIGKLC14	Synthetic c
7	66	100.0	711	12	HSIGKLC31	Synthetic c
8	66	100.0	1128	6	AX431268	Sequence
9	66	100.0	1128	6	AX431271	Sequence
10	66	100.0	3394	12	AF268280	Phagemid
11	66	100.0	3450	6	AR241678	Sequence
12	66	100.0	3758	12	AF268281	Phagemid
13	66	100.0	5149	6	AX591634	Sequence
14	66	100.0	5149	6	AX600065	Sequence
15	66	100.0	5683	6	AX554424	Sequence
16	66	100.0	6122	6	AX554413	Sequence
17	66	100.0	6122	6	AX554422	Sequence
18	66	100.0	6122	6	AX642149	Sequence
19	66	100.0	6125	6	AX554420	Sequence
20	64.4	97.6	5885	1	STAFRA	L42764 Staphylococ
21	61.2	92.7	580	6	AX742885	Sequence
22	61.2	92.7	580	6	AX743509	Sequence
23	60.2	91.2	74	6	E02814	E02814 DNA sequenc
24	60.2	91.2	74	6	E03955	E03955 DNA encodin
25	60.2	91.2	103	6	AX236721	Sequence
26	60.2	91.2	249	6	I08200	Sequence 1
27	60.2	91.2	249	6	I52035	Sequence 1
28	60.2	91.2	438	12	SYNOMPANA	M62710 Synthetic g
29	60.2	91.2	548	6	ARI79433	Sequence
30	60.2	91.2	548	6	AR282580	Sequence
31	60.2	91.2	548	6	AX022441	Sequence
32	60.2	91.2	548	6	AX030798	Sequence
33	60.2	91.2	548	6	AX034619	Sequence
34	60.2	91.2	652	6	E01685	E01685 hybrid DNA
35	60.2	91.2	739	1	AF234271	AF234271 Shigella
36	60.2	91.2	746	1	AF234270	AF234270 Salmonell
37	60.2	91.2	1400	1	STOMPA	X02006 Salmonella
38	60.2	91.2	1697	1	AF234269	AF234269 Escherich
39	60.2	91.2	2270	6	AX191724	Sequence
40	60.2	91.2	2271	1	E00458	V00307 E. coli omp
41	60.2	91.2	6477	6	A08458	A08458 pinf 4-49 m
42	60.2	91.2	7499	12	ASAK3121	AJ223121 Artificia
43	60.2	91.2	7500	12	ASAK3122	AJ223122 Artificia
c 44	60.2	91.2	11639	1	AE000198	AE000198 Escherich
c 45	60.2	91.2	11641	1	AE005286	AE005286 Escherich

ALIGNMENTS

RESULT 1  
AX430829  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL

AX430829  
Sequence 1 from Patent WO0240696.  
AX430829  
GI:21655908  
Escherichia coli  
Escherichia coli  
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
Enterobacteriaceae; Escherichia.  
1  
Goetz, F., Werner, R.G., Tayapiwatana, C., Manosroi, A. and Manosroi, J.  
Methods for large scale protein production in prokaryotes  
Patent: WO 0240696-A 1 23-MAY-2002;

AX430829  
Sequence 1 from Patent WO0240696.  
AX430829  
GI:21655908  
Escherichia coli  
Escherichia coli  
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
Enterobacteriaceae; Escherichia.  
1  
Goetz, F., Werner, R.G., Tayapiwatana, C., Manosroi, A. and Manosroi, J.  
Methods for large scale protein production in prokaryotes  
Patent: WO 0240696-A 1 23-MAY-2002;

AX430829  
Sequence 1 from Patent WO0240696.  
AX430829  
GI:21655908  
Escherichia coli  
Escherichia coli  
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
Enterobacteriaceae; Escherichia.  
1  
Goetz, F., Werner, R.G., Tayapiwatana, C., Manosroi, A. and Manosroi, J.  
Methods for large scale protein production in prokaryotes  
Patent: WO 0240696-A 1 23-MAY-2002;

```

BOEHRINGER INGELHEIM INT (DE)
FEATURES
  source
    1. .66
    /organism="Escherichia coli"
    /mol_type="genomic DNA"
    /db_xref="taxon:562"
  BASE COUNT      14 a      21 g      13 t
  ORIGIN
    Query Match      100.0%; Score 66; DB 6; Length 66;
    Best Local Similarity 100.0%; Pred. No. 1.4e-11;
    Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
  QY      1 ATGAAAAGACAGCTATCGGATTGCGACATGGCGCTGTTTCGCTACCGTGCCCGAG 60
  DB      1 ATGAAAAGACAGCTATCGGATTGCGACATGGCGCTGTTTCGCTACCGTGCCCGAG 60
  QY      61 GCGGCC 66
  DB      61 GCGGCC 66

RESULT 2
AX431269
LOCUS      AX431269      66 bp      DNA      linear      PAT 28-JUN-2002
DEFINITION      Sequence 3 from Patent WO0240650.
ACCESSION      AX431269
VERSION      AX431269.1 GI:21656151
KEYWORDS
  SOURCE
    Escherichia coli
    Escherichia coli
    Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
    Enterobacteriaceae; Escherichia.
  REFERENCE
    1 Goetz,F., Werner,R.G., Tayapiwatana,C., Manosroi,A. and Manosroi,J.
    METHODS for large scale production of recombinant dna-derived tpa
    or k2s molecules
    Patent: WO 0240650-A 3 23-MAY-2002;
  JOURNAL      BOEHRINGER INGELHEIM INT (DE)
  FEATURES
    source
      1. .66
      /organism="Escherichia coli"
      /mol_type="genomic DNA"
      /db_xref="taxon:562"
    BASE COUNT      14 a      21 g      13 t
    ORIGIN
      Query Match      100.0%; Score 66; DB 6; Length 66;
      Best Local Similarity 100.0%; Pred. No. 1.4e-11;
      Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
    QY      1 ATGAAAAGACAGCTATCGGATTGCGACATGGCGCTGTTTCGCTACCGTGCCCGAG 60
    DB      1 ATGAAAAGACAGCTATCGGATTGCGACATGGCGCTGTTTCGCTACCGTGCCCGAG 60
    QY      61 GCGGCC 66
    DB      61 GCGGCC 66

RESULT 3
AX431272
LOCUS      AX431272      66 bp      DNA      linear      PAT 28-JUN-2002
DEFINITION      Sequence 6 from Patent WO0240650.
ACCESSION      AX431272
VERSION      AX431272.1 GI:21656154
KEYWORDS
  SOURCE
    Escherichia coli
    Escherichia coli
    Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
    Enterobacteriaceae; Escherichia.
  REFERENCE
    1 Goetz,F., Werner,R.G., Tayapiwatana,C., Manosroi,A. and Manosroi,J.
    METHODS for large scale production of recombinant dna-derived tpa
    or k2s molecules
    Patent: WO 0240650-A 3 23-MAY-2002;
  JOURNAL      BOEHRINGER INGELHEIM INT (DE)
  FEATURES
    source
      1. .66
      /organism="Escherichia coli"
      /mol_type="genomic DNA"
      /db_xref="taxon:562"
    BASE COUNT      14 a      21 g      13 t
    ORIGIN
      Query Match      100.0%; Score 66; DB 6; Length 66;
      Best Local Similarity 100.0%; Pred. No. 1.4e-11;
      Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
    QY      1 ATGAAAAGACAGCTATCGGATTGCGACATGGCGCTGTTTCGCTACCGTGCCCGAG 60
    DB      1 ATGAAAAGACAGCTATCGGATTGCGACATGGCGCTGTTTCGCTACCGTGCCCGAG 60
    QY      61 GCGGCC 66
    DB      61 GCGGCC 66

RESULT 4
HSIGKLC28
LOCUS      HSIGKLC28      705 bp      mRNA      linear      SYN 29-MAR-2001
DEFINITION      Synthetic construct including Homo sapiens immunoglobulin kappa
chain (clone: 28).
ACCESSION      X95748
VERSION      X95748.1 GI:1514578
KEYWORDS
  SOURCE
    constant region; immunoglobulin; kappa light chain.
    synthetic construct
    synthetic construct
    artificial sequences.
  REFERENCE
    1 Steinberger,P., Kraft,D. and Valenta,R.
    CONSTRUCTION of a combinatorial IGE library from an allergic
    patient. Isolation and characterization of human IGE Fabs with
    specificity for the major timothy grass pollen allergen, Phi p 5
    J. Biol. Chem. 271 (18), 10967-10972 (1996)
  JOURNAL      J. Biol. Chem. 271 (18), 10967-10972 (1996)
  MEDLINE      96210038
  PUBMED      8631916
  REFERENCE
    2 (bases 1 to 705)
  AUTHORS
    Valenta,R.L.S.
  TITLE
    Direct Submission
  JOURNAL
    Submitted (19-FEB-1996) R.L.S. Valenta, Institute of General &
    Experimental Pathology, General Hospital, Waehringer Guertel 18-20,
    1090 Vienna, AUSTRIA
  FEATURES
    Location/Qualifiers
      1. .705
      /organism="synthetic construct"
      /mol_type="mRNA"
      /db_xref="taxon:32630"
      /focus
      1. .66
      /organism="Escherichia coli"
      /mol_type="mRNA"
      /db_xref="taxon:562"
      67. .705
      /organism="Homo sapiens"
      /mol_type="mRNA"
      /db_xref="taxon:9606"
      /clone="28"
      /cell_type="B-cell"
      /clone_lib="pcomb3H"
      1. .705
      /genes="OMP-A"
      1. .705
      /genes="OMP-A"
      /codon_start=1
      /transl_table=11

```



```

/gene="OMP-A"
/note="FR4"
385..702
/gene="OMP-A"
/note="kappa constant region"
661..705
/gene="OMP-A"
/note="PCR-primer sequence"
183 a 191 c 181 g 150 t

BASE COUNT 183 a 191 c 181 g 150 t
ORIGIN

Query Match 100.0%; Score 66; DB 12; Length 705;
Best Local Similarity 100.0%; Pred. No. 1.6e-11;
Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGAAAAGACAGCTATCGGATTCAGTGGCAGTGGCTGGTTTCGCTACCGTGCCCGCAG 60
Db 1 ATGAAAAGACAGCTATCGGATTCAGTGGCAGTGGCTGGTTTCGCTACCGTGCCCGCAG 60

Qy 61 GCGGCC 66
Db 61 GCGGCC 66

RESULT 6
HSIGKLC14 711 bp mRNA linear SYN 29-MAR-2001
LOCUS Synthetic construct including Homo sapiens immunoglobulin kappa
DEFINITION chain (clone: 14).
ACCESSION X95749
VERSION X95749.1 GI:1514576
KEYWORDS constant region; immunoglobulin; kappa light chain.
SOURCE synthetic construct
ORGANISM synthetic construct
artificial sequences.
REFERENCE 1
AUTHORS Steinberger,P., Kraft,D. and Valenta,R.
TITLE Construction of a combinatorial IgE library from an allergic
patient. Isolation and characterization of human IgE Fabs with
specificity for the major timothy grass pollen allergen, Phl p 5
J. Biol. Chem. 271 (18), 10967-10972 (1996)
PUBMED 96210038
REFERENCE 2 (bases 1 to 711)
AUTHORS Valenta,R.L.S.
TITLE Direct Submission
JOURNAL Submitted (19-FEB-1996) R.L.S. Valenta, Institute of General &
Experimental Pathology, General Hospital, Waehringer Guertel 18-20,
1090 Vienna, AUSTRIA
FEATURES
source 1..711
/organism="synthetic construct"
/mol_type="mRNA"
/db_xref="taxon:32630"
/focus
source 1..66
/organism="Escherichia coli"
/mol_type="mRNA"
/db_xref="taxon:562"
source 67..711
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="14"
/cell_type="B-cell"
/clone_lib="pcmb3H"
1..711
/gene="OMP-A"
1..711
/gene="OMP-A"
/codon_start=1
/transl_table=11
/product="immunoglobulin kappa light chain"

/protein_id="CAA65060.1"
/db_xref="GI:1514577"
/db_xref="REMBL:CAA65060"
/translation="MKKTAIAIALAGFATVAQAELVMTQSPSSLSASVGDRTVIT
CRASHYSNLYNMWYQQKPKAPKLLIYAASLQSGVPSRFGSGDFTLTISLQ
EDFANYQCQSFSGTGGTQVEIKRTVAAPSVFIFPSDEQLKSGTASVCLINN
FYPREAKVQWKNALQSGNSQESVTEQDSKDSSTLSLTLSKADYKHKLYACEV
THQGLSPVTKSFNRGEC"
1..66
/gene="OMP-A"
/note="leader peptide gene"
67..135
/gene="OMP-A"
/note="FR1"
67..84
/gene="OMP-A"
/note="PCR-primer sequence"
67..72
/gene="OMP-A"
/note="SAC 1 restriction site"
136..168
/gene="OMP-A"
/note="CDR1"
169..213
/gene="OMP-A"
/note="FR2"
214..234
/gene="OMP-A"
/note="CDR2"
235..330
/gene="OMP-A"
/note="FR3"
331..357
/gene="OMP-A"
/note="CDR3"
358..390
/gene="OMP-A"
/note="FR4"
391..708
/gene="OMP-A"
/note="kappa constant region"
667..708
/gene="OMP-A"
/note="PCR-primer sequence"
187 a 194 c 179 g 151 t

BASE COUNT 187 a 194 c 179 g 151 t
ORIGIN

Query Match 100.0%; Score 66; DB 12; Length 711;
Best Local Similarity 100.0%; Pred. No. 1.6e-11;
Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGAAAAGACAGCTATCGGATTCAGTGGCAGTGGCTGGTTTCGCTACCGTGCCCGCAG 60
Db 1 ATGAAAAGACAGCTATCGGATTCAGTGGCAGTGGCTGGTTTCGCTACCGTGCCCGCAG 60

Qy 61 GCGGCC 66
Db 61 GCGGCC 66

RESULT 7
HSIGKLC31 711 bp mRNA linear SYN 29-MAR-2001
LOCUS Synthetic construct including Homo sapiens immunoglobulin kappa
DEFINITION chain (clone: 31).
ACCESSION X95750
VERSION X95750.1 GI:1514582
KEYWORDS constant region; immunoglobulin; kappa light chain.
SOURCE synthetic construct
ORGANISM synthetic construct
artificial sequences.
REFERENCE 1
AUTHORS Steinberger,P., Kraft,D. and Valenta,R.

```

TITLE Construction of a combinatorial IgE library from an allergic patient. Isolation and characterization of human IgE Fabs with specificity for the major timothy grass pollen allergen, Phl p 5

JOURNAL J. Biol. Chem. 271 (18), 10967-10972 (1996)

MEDLINE 96210038

PUBMED 8631916

REFERENCE 2 (bases 1 to 711)

AUTHORS Valenta, R.L.S.

TITLE Direct Submission

JOURNAL Submitted (19-FEB-1996) R.L.S. Valenta, Institute of General & Experimental Pathology, General Hospital, Waehringer Guertel 18-20, 1090 Vienna, AUSTRIA

FEATURES

source

1..711  
 Location/Qualifiers  
 /organism="synthetic construct"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:32630"  
 /focus

source

1..66  
 /organism="Escherichia coli"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:562"  
 67..711  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="31"  
 /cell\_type="B-cell"  
 /clone\_lib="pcomb3H"  
 1..711  
 /gene="OMP-A"  
 1..711  
 /gene="OMP-A"  
 /codon\_start=1  
 /transl\_table=11  
 /product="immunoglobulin kappa light chain"  
 /protein\_id="CAA65061.1"  
 /db\_xref="GI:1514583"  
 /db\_xref="REMBL:CAA65061"  
 /translation="MKKTAIAIAVALAGFATVAOAEIIVMTQSPSSLSAFVGDRTVITCRASQSISSLYNNWQOKPKAPKLLIYAASSLQSGVPSRFGSGGTDFTLTISLQPEDFATYICQSQSHSPYTPGQTNLEIKRTVAAPSVFIFPPSDEQLKSGTASVCLLNNFYPRKAVQWKNALQSGNSQESVTEQDSKDSSTYSSTLTLSKADYERKHVACEVTHQGLSSPVTKSFNRGEC"

mrna

1..66  
 /gene="OMP-A"  
 /note="leaderpeptide gene"

misc\_feature

67..135  
 /gene="OMP-A"  
 /note="FR1"

primer\_bind

67..84  
 /gene="OMP-A"  
 /note="PCR-primer sequence"

misc\_feature

67..72  
 /gene="OMP-A"  
 /note="SAC 1 restriction site"

misc\_feature

136..168  
 /gene="OMP-A"  
 /note="CDR1"

misc\_feature

169..213  
 /gene="OMP-A"  
 /note="FR2"

misc\_feature

214..234  
 /gene="OMP-A"  
 /note="CDR2"

misc\_feature

235..330  
 /gene="OMP-A"  
 /note="FR3"

misc\_feature

331..357  
 /gene="OMP-A"  
 /note="CDR3"

misc\_feature

358..390  
 /gene="OMP-A"

C\_region

primer\_bind

BASE COUNT 188 a 194 c 179 g 150 t  
 ORIGIN

Query Match 100.0%; Score 66; DB 12; Length 711;  
 Best Local Similarity 100.0%; Pred. No. 1.6e-11;  
 Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGAAAAGACAGCTATCGGATTGCGACTGCGTGGTTCGCTACCGTGCCCGAG 60  
 |||||  
 Db 1 ATGAAAAGACAGCTATCGGATTGCGACTGCGTGGTTCGCTACCGTGCCCGAG 60  
 |||||

Qy 61 GCGGCC 66  
 |||||  
 Db 61 GCGGCC 66

RESULT 8  
 AX431268  
 LOCUS AX431268 1128 bp DNA linear PAT 28-JUN-2002  
 DEFINITION Sequence 2 from Patent WO0240650.  
 ACCESSION AX431268  
 VERSION AX431268.1 GI:21656150  
 KEYWORDS  
 SOURCE synthetic construct  
 ORGANISM synthetic construct  
 REFERENCE 1  
 AUTHORS Goetz, F., Werner, R.G., Tayapiwatana, C., Manosroi, A. and Manosroi, J.  
 TITLE Methods for large scale production of recombinant dna-derived tpa or k2s molecules  
 JOURNAL Patent: WO 0240650-A 2 23-MAY-2002;  
 BOHRINGER INGELHEIM INT (DE)  
 FEATURES  
 source  
 1..1128  
 /organism="synthetic construct"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:32630"  
 /note="coding sequence for OmpA-K2S fusion protein"

BASE COUNT 238 a 332 c 332 g 226 t  
 ORIGIN

Query Match 100.0%; Score 66; DB 6; Length 1128;  
 Best Local Similarity 100.0%; Pred. No. 1.6e-11;  
 Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGAAAAGACAGCTATCGGATTGCGACTGCGTGGTTCGCTACCGTGCCCGAG 60  
 |||||  
 Db 1 ATGAAAAGACAGCTATCGGATTGCGACTGCGTGGTTCGCTACCGTGCCCGAG 60  
 |||||

Qy 61 GCGGCC 66  
 |||||  
 Db 61 GCGGCC 66

RESULT 9  
 AX431271  
 LOCUS AX431271 1128 bp DNA linear PAT 28-JUN-2002  
 DEFINITION Sequence 5 from Patent WO0240650.  
 ACCESSION AX431271  
 VERSION AX431271.1 GI:21656153  
 KEYWORDS  
 SOURCE synthetic construct  
 ORGANISM synthetic construct  
 REFERENCE 1  
 AUTHORS Goetz, F., Werner, R.G., Tayapiwatana, C., Manosroi, A. and Manosroi, J.

TITLE Methods for large scale production of recombinant dna-derived tpa or k2s molecules  
JOURNAL Patent: WO 0240650-A 5 23-MAY-2002;  
BOEHRINGER INGELHEIM INT (DE)  
FEATURES Location/Qualifiers  
source 1..1128  
/organism="synthetic construct"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:32630"  
/note="coding sequence for OmpA-K2S fusion protein"  
BASE COUNT 238 a 332 c 332 g 226 t  
ORIGIN  
Query Match 100.0%; Score 66; DB 6; Length 1128;  
Best Local Similarity 100.0%; Pred. No. 1.7e-11;  
Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 ATGAAAAGACAGCTATCGGATTCAGTGGCACTGGCTGGTTTCGCTACCGTGCCCGAG 60  
|||||  
Db 1 ATGAAAAGACAGCTATCGGATTCAGTGGCACTGGCTGGTTTCGCTACCGTGCCCGAG 60  
|||||  
Qy 61 GCGGCC 66  
|||||  
Db 61 GCGGCC 66  
|||||  
RESULT 10  
AF268280 3394 bp DNA circular SYN 03-OCT-2000  
LOCUS Phagemid cloning vector pComb3H, complete sequence.  
DEFINITION AF268280.1 GI:10505048  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE Phagemid cloning vector pComb3H  
ORGANISM Phagemid cloning vector pComb3H  
Phagemid cloning vector pComb3H  
artificial sequences; vectors.  
REFERENCE 1 (bases 1 to 3394)  
AUTHORS Rader,C. and Barbas,C.F. III.  
JOURNAL (in) PHAGE DISPLAY, A LABORATORY MANUAL. Cold Spring Harbor  
Laboratory Press, Cold Spring Harbor, NY, USA (2000), In press  
REFERENCE 2 (bases 1 to 3394)  
AUTHORS Rader,C. and Barbas,C.F. III.  
JOURNAL Direct Submission  
TITLE Submitted (16-MAY-2000) Department of Molecular Biology, BCC-526,  
Scripps Research Institute, 10550 N. Torrey Pines Rd., La Jolla, CA  
92037, USA  
FEATURES Location/Qualifiers  
source 1..3394  
/organism="Phagemid cloning vector pComb3H"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:137785"  
BASE COUNT 862 a 783 c 876 g 873 t  
ORIGIN  
Query Match 100.0%; Score 66; DB 12; Length 3394;  
Best Local Similarity 100.0%; Pred. No. 1.7e-11;  
Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 ATGAAAAGACAGCTATCGGATTCAGTGGCACTGGCTGGTTTCGCTACCGTGCCCGAG 60  
|||||  
Db 222 ATGAAAAGACAGCTATCGGATTCAGTGGCACTGGCTGGTTTCGCTACCGTGCCCGAG 281  
|||||  
Qy 61 GCGGCC 66  
|||||  
Db 282 GCGGCC 287  
|||||  
RESULT 11  
AR241678 3450 bp DNA linear PAT 20-DEC-2002  
LOCUS Sequence 20 from patent US 6472147.  
DEFINITION AR241678  
ACCESSION AR241678  
VERSION AR241678.1 GI:27287480

KEYWORDS Unknown.  
SOURCE Unknown.  
ORGANISM Unclassified.  
REFERENCE 1 (bases 1 to 3450)  
AUTHORS Janda,K.D., Wirsching,P., Lerner,R.A. and Gao,C.  
TITLE Methods for display of heterodimeric proteins on filamentous phage  
using pVII and pIX, compositions, vectors and combinatorial  
libraries  
JOURNAL Patent: US 6472147-A 20 29-OCT-2002;  
FEATURES Location/Qualifiers  
source 1..3450  
/organism="unknown"  
BASE COUNT 873 a 799 c 891 g 887 t  
ORIGIN  
Query Match 100.0%; Score 66; DB 6; Length 3450;  
Best Local Similarity 100.0%; Pred. No. 1.7e-11;  
Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 ATGAAAAGACAGCTATCGGATTCAGTGGCACTGGCTGGTTTCGCTACCGTGCCCGAG 60  
|||||  
Db 2667 ATGAAAAGACAGCTATCGGATTCAGTGGCACTGGCTGGTTTCGCTACCGTGCCCGAG 2726  
|||||  
Qy 61 GCGGCC 66  
|||||  
Db 2727 GCGGCC 2732  
|||||  
RESULT 12  
AF268281 3758 bp DNA circular SYN 03-OCT-2000  
LOCUS Phagemid cloning vector pComb3X, complete sequence.  
DEFINITION AF268281  
ACCESSION AF268281  
VERSION AF268281.1 GI:10505049  
KEYWORDS  
SOURCE Phagemid cloning vector pComb3X  
ORGANISM Phagemid cloning vector pComb3X  
Phagemid cloning vector pComb3X  
artificial sequences; vectors.  
REFERENCE 1 (bases 1 to 3758)  
AUTHORS Rader,C. and Barbas,C.F. III.  
JOURNAL (in) PHAGE DISPLAY, A LABORATORY MANUAL. Cold Spring Harbor  
Laboratory Press, Cold Spring Harbor, NY, USA (2000), In press  
REFERENCE 2 (bases 1 to 3758)  
AUTHORS Rader,C. and Barbas,C.F. III.  
JOURNAL Direct Submission  
TITLE Submitted (16-MAY-2000) Department of Molecular Biology, BCC-526,  
Scripps Research Institute, 10550 N. Torrey Pines Rd., La Jolla, CA  
92037, USA  
FEATURES Location/Qualifiers  
source 1..3758  
/organism="Phagemid cloning vector pComb3X"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:137786"  
BASE COUNT 939 a 896 c 959 g 964 t  
ORIGIN  
Query Match 100.0%; Score 66; DB 12; Length 3758;  
Best Local Similarity 100.0%; Pred. No. 1.7e-11;  
Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 ATGAAAAGACAGCTATCGGATTCAGTGGCACTGGCTGGTTTCGCTACCGTGCCCGAG 60  
|||||  
Db 222 ATGAAAAGACAGCTATCGGATTCAGTGGCACTGGCTGGTTTCGCTACCGTGCCCGAG 281  
|||||  
Qy 61 GCGGCC 66  
|||||  
Db 282 GCGGCC 287  
|||||  
RESULT 13  
AX591634 5149 bp DNA linear PAT 27-JAN-2003  
LOCUS AX591634



```

DEFINITION Sequence 8 from Patent WO0246434.
ACCESSION AX591634
VERSION AX591634.1 GI:27950030
KEYWORDS
SOURCE
ORGANISM
synthetic construct
artificial sequences.
REFERENCE
1
AUTHORS Bowdish,K.S. and Barbas-Frederickson,S.
TITLE Novel plasmid vectors
JOURNAL Patent: WO 0246434-A 13-JUN-2002;
ALEXION PHARMACEUTICALS, INC. (US)
FEATURES
Location/Qualifiers
source
1..5149
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
/note="vector"
BASE COUNT 1279 a 1244 c 1372 g 1254 t
ORIGIN
Query Match 100.0%; Score 66; DB 6; Length 5149;
Best Local Similarity 100.0%; Pred. No. 1.8e-11;
Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGAAAAGACAGCTATCGGATTCAGTGGCAGTGGCTGGTTTCGCTACCGTGCCCGAG 60
DB 2611 ATGAAAAGACAGCTATCGGATTCAGTGGCAGTGGCTGGTTTCGCTACCGTGCCCGAG 2670
QY 61 GCGGCC 66
DB 2671 GCGGCC 2676

RESULT 14
AX600065
LOCUS AX600065 5149 bp DNA linear PAT 14-FEB-2003
DEFINITION Sequence 60 from Patent WO0246238.
ACCESSION AX600065
VERSION AX600065.1 GI:28400138
KEYWORDS
SOURCE
ORGANISM
synthetic construct
artificial sequences.
REFERENCE
1
AUTHORS Bowdish,K.S., Barbas-Frederickson,S. and Renshaw,M.
TITLE Rationally designed antibodies
JOURNAL Patent: WO 0246238-A 60 13-JUN-2002;
ALEXION PHARMACEUTICALS, INC. (US)
FEATURES
Location/Qualifiers
source
1..5149
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
/note="vector"
BASE COUNT 1279 a 1244 c 1372 g 1254 t
ORIGIN
Query Match 100.0%; Score 66; DB 6; Length 5149;
Best Local Similarity 100.0%; Pred. No. 1.8e-11;
Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGAAAAGACAGCTATCGGATTCAGTGGCAGTGGCTGGTTTCGCTACCGTGCCCGAG 60
DB 2611 ATGAAAAGACAGCTATCGGATTCAGTGGCAGTGGCTGGTTTCGCTACCGTGCCCGAG 2670
QY 61 GCGGCC 66
DB 2671 GCGGCC 2676

RESULT 15
AX554424
LOCUS AX554424 5683 bp DNA linear PAT 27-NOV-2002
DEFINITION Sequence 12 from Patent WO0246436.
ACCESSION AX554424
VERSION AX554424.1 GI:25898200
KEYWORDS
SOURCE
ORGANISM
synthetic construct
artificial sequences.
REFERENCE
1
AUTHORS Bowdish,K.S., Barbas-Frederickson,S., Wild,M. and McWhirter,J.
TITLE Novel plasmid vectors
JOURNAL Patent: WO 0246436-A 12 13-JUN-2002;
ALEXION PHARMACEUTICALS, INC. (US)
FEATURES
Location/Qualifiers
source
1..5683
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
/note="plasmid"
BASE COUNT 1337 a 1492 c 1543 g 1311 t
ORIGIN
Query Match 100.0%; Score 66; DB 6; Length 5683;
Best Local Similarity 100.0%; Pred. No. 1.8e-11;
Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGAAAAGACAGCTATCGGATTCAGTGGCAGTGGCTGGTTTCGCTACCGTGCCCGAG 60
DB 2611 ATGAAAAGACAGCTATCGGATTCAGTGGCAGTGGCTGGTTTCGCTACCGTGCCCGAG 2670
QY 61 GCGGCC 66
DB 2671 GCGGCC 2676

Search completed: August 19, 2003, 04:46:31
Job time : 281.124 secs
```

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 18, 2003, 19:51:53 ; Search time 23.2513 Seconds  
(without alignments)  
7662.485 Million cell updates/sec

Title: US-09-987-455-3  
Perfect score: 66  
Sequence: 1 atgaaaagacagtatcg.....ctacgtggccagcgagcc 66

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2582756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Maximum Match 0%  
Listing first 45 summaries

Database : N\_Geneseq\_19Jun03.\*  
1: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.\*  
2: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.\*  
3: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.\*  
4: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.\*  
5: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.\*  
6: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.\*  
7: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.\*  
8: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.\*  
9: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.\*  
10: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.\*  
11: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.\*  
12: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.\*  
13: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.\*  
14: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.\*  
15: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.\*  
16: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.\*  
17: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.\*  
18: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.\*  
19: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.\*  
20: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.\*  
21: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.\*  
22: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.\*  
23: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.\*  
24: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.\*  
25: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	66	100.0	66	AAD40615	Escherichia coli o
2	66	100.0	66	AAD38987	Escherichia coli O
3	66	100.0	1128	AAD40614	OmpA-K2S fusion pr
4	66	100.0	1574	AAZ57599	Erythrocyte and He
5	66	100.0	3450	AAZ66993	Vector pCGMT-1b.
6	66	100.0	4691	AAQ92546	pCom3 expression
7	66	100.0	5149	ABQ73372	Plasmid pRL8 nucle
8	66	100.0	5149	ABN84077	Phagemid vector pr

9	66	100.0	5683	24	AA411116	Plasmid pRL5 nucle
10	66	100.0	6122	24	AA411112	Plasmid pRL5 CAT n
11	66	100.0	6122	24	AA411115	Plasmid pRL5 bsi-C
12	66	100.0	6122	24	ABN84078	Plasmid pRL5-CAT.
13	66	100.0	6125	24	AA411114	Plasmid pRL5 CAT-A
14	66	100.0	6166	16	AAQ92547	Expression vector,
15	60.2	91.2	73	12	AAQ10569	Omp A signal pepti
16	60.2	91.2	103	22	AAQ44154	Neisseria meningit
17	60.2	91.2	249	11	AAQ03521	TAC recombinant DN
18	60.2	91.2	548	18	AAQ91586	Growth factor LHL
19	60.2	91.2	548	20	AAQ34590	DNA encoding a LHL
20	60.2	91.2	652	9	AAQ81344	Sequence encoding
21	60.2	91.2	2814	14	AAQ41731	PE binding/translo
22	60.2	91.2	2814	14	AAQ38414	PE binding and tra
23	60.2	91.2	6477	11	AAQ02030	Plasmid pinf 4-49.
24	59.6	90.3	73	21	AA412579	Synthetic DNA enco
25	59.6	90.3	415	17	AAQ15982	psKAN8 fragment #1
26	59.6	90.3	470	18	AAQ91590	Kappa light chain
27	59.6	90.3	470	20	AAQ34594	DNA encoding a hum
28	59.6	90.3	599	18	AAQ91588	Growth factor LHL.
29	59.6	90.3	599	20	AAQ34593	DNA encoding a for
30	59.6	90.3	759	19	AAQ53721	The ORF of the clo
31	59.6	90.3	770	19	AAQ53720	Cloned duplicate o
32	59.6	90.3	793	20	AAQ24101	Plasmid pBBP22 DNA
33	59.6	90.3	793	22	AAQ25703	Bilin binding-prot
34	59.6	90.3	803	12	AAQ14831	Qm212 single chain
35	59.6	90.3	932	22	AAQ91151	Expression cassett
36	59.6	90.3	1031	18	AAQ91589	Growth factor TLHL
37	59.6	90.3	1031	20	AAQ34592	DNA encoding TLHL.
38	59.6	90.3	1219	20	AAQ24099	Plasmid pBBP20 DNA
39	59.6	90.3	1219	22	AAQ25695	Bilin binding-prot
40	59.6	90.3	1219	22	AAQ25708	Bilin binding-prot
41	59.6	90.3	1380	20	AAQ24100	Plasmid pBBP21 DNA
42	59.6	90.3	1380	22	AAQ25710	Bilin binding-prot
43	59.6	90.3	1479	20	AAQ34596	DNA encoding a ccm
44	59.6	90.3	1490	18	AAQ91587	Growth factor CATA
45	59.6	90.3	1490	20	AAQ34591	DNA encoding CATAB

ALIGNMENTS

RESULT 1	
AAD40615	
ID	AAD40615 standard; DNA; 66 BP.
XX	AAD40615;
AC	AAD40615;
XX	30-OCT-2002 (first entry)
DT	Escherichia coli ompA peptide encoding DNA.
XX	Tissue plasminogen activator; tPA; kringle 2 serine protease; stroke;
KW	K2S; cardiac infarction; acute myocardial infarction; pulmonary embolism;
KW	artery occlusion; deep vein thrombosis; blood clotting; thrombolytic;
KW	cerebroprotective; cardiant; ompA; gene; ds.
XX	Escherichia coli.
OS	Escherichia coli.
XX	Key
FH	Location/Qualifiers
FT	1..66
FT	/*tag= a
FT	/product= "OmpA peptide"
FT	/note= "No stop codon"
FT	/partial
XX	
XX	WO200240650-A2.
FN	
XX	
PD	23-MAY-2002.
XX	
PF	07-NOV-2001; 2001WO-EPI2857.
XX	
PR	14-NOV-2000; 2000GB-0027779.

XX (BOEH ) BOEHRINGER INGELHEIM INT GMBH.  
 XX Werner R, Goetz F, Tayapiwatana C, Manosroi J, Manosroi A;  
 XX WPI; 2002-519376/55.  
 XX P-PSDB; AAE25035.  
 XX  
 XX Producing active, correctly folded recombinant tissue plasminogen  
 PT activator, Kringle 2 serine protease in prokaryotic cells by expressing  
 PT the protein-encoding DNA operably linked to DNA coding for signal  
 PT peptide OmpA  
 XX  
 XX Claim 8; Page 30; 80pp; English.  
 XX  
 XX The present invention relates to a method of producing extracellularly  
 CC secreted, active, correctly folded, recombinant tissue plasminogen  
 CC activator (tPA), Kringle 2 serine protease molecule (K2S) or their  
 CC variants in prokaryotic cells by expressing the protein-encoding DNA  
 CC operably linked to DNA coding for signal peptide OmpA. The method is  
 CC useful for producing recombinant DNA-derived tPA, K2S or their variants.  
 CC Sequences of the invention are useful for manufacturing a medicament  
 CC for treating stroke, cardiac infarction, acute myocardial infarction,  
 CC pulmonary embolism, any artery occlusion such as intracranial artery  
 CC occlusion (e.g., arteries supplying the brain), peripherally occluded  
 CC arteries, coronary artery occlusion, deep vein thrombosis or related  
 CC diseases associated with unwanted blood clotting. The present sequence  
 CC is a DNA fragment encoding Escherichia coli OmpA peptide.  
 XX  
 XX Sequence 66 BP; 14 A; 18 C; 21 G; 13 T; 0 other;  
 SQ  
 Query Match 100.0%; Score 66; DB 24; Length 66;  
 Best Local Similarity 100.0%; Pred. No. 2.1e-14;  
 Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 ATGAAAAGACAGCTATCGCGATTGCGAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCCAG 60  
 DB 1 ATGAAAAGACAGCTATCGCGATTGCGAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCCAG 60  
 QY 61 GCGGCC 66  
 DB 61 GCGGCC 66  
 RESULT 2  
 AAD38987  
 ID AAD38987 standard; DNA; 66 BP.  
 AC AAD38987;  
 XX  
 XX 23-SEP-2002 (first entry)  
 XX  
 XX Escherichia coli OmpA DNA.  
 DE  
 XX Recombinant DNA-derived heterologous protein; OmpA; K2S; kringle 2;  
 KW serine protease; ds.  
 KW  
 XX Escherichia coli.  
 OS  
 XX WO200240696-A2.  
 XX  
 XX 23-MAY-2002.  
 XX  
 XX 08-NOV-2001; 2001WO-EP12920.  
 XX  
 XX 14-NOV-2000; 2000GB-0027782.  
 XX  
 XX (BOEH ) BOEHRINGER INGELHEIM INT GMBH.  
 XX  
 XX Werner R, Goetz F, Tayapiwatana C, Manosroi J, Manosroi A;  
 XX WPI; 2002-471625/50.  
 XX  
 XX  
 PT Producing recombinant DNA-derived kringle 2 plus serine protease,  
 PT comprises using a prokaryotic cell expressing a vector having a DNA  
 PT coding for a heterologous protein operably linked to a DNA coding for  
 PT the signal peptide OmpA  
 XX  
 XX Claim 9; Page 23; 52pp; English.  
 XX  
 XX The invention relates to a method for producing recombinant DNA-derived  
 CC heterologous protein in prokaryotic cells, where the heterologous protein  
 CC is secreted extracellularly as an active and correctly folded protein and  
 CC the prokaryotic cell contains and expresses a vector comprising the DNA  
 CC coding for the heterologous protein operably linked to the DNA coding  
 CC for the signal peptide OmpA or its functional derivative. The method is  
 CC useful for commercial large-scale production of heterologous proteins,  
 CC e.g. K2S (kringle 2 plus serine protease), in prokaryotic cells, and is  
 CC generally applicable in the expression of several different proteins and  
 CC polypeptides which do not require mammalian glycosylation in prokaryotic  
 CC host cells. The method may also be used to obtain DNA sequences of a  
 CC protein of interest to be expressed from databases and cloned for use.  
 CC The present sequence is Escherichia coli OmpA DNA.  
 XX  
 XX Sequence 66 BP; 14 A; 18 C; 21 G; 13 T; 0 other;  
 SQ  
 Query Match 100.0%; Score 66; DB 24; Length 66;  
 Best Local Similarity 100.0%; Pred. No. 2.1e-14;  
 Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 ATGAAAAGACAGCTATCGCGATTGCGAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCCAG 60  
 DB 1 ATGAAAAGACAGCTATCGCGATTGCGAGTGGCACTGGCTGGTTTCGCTACCGTGGCCCCAG 60  
 QY 61 GCGGCC 66  
 DB 61 GCGGCC 66  
 RESULT 3  
 AAD40614  
 ID AAD40614 standard; DNA; 1128 BP.  
 AC AAD40614;  
 XX  
 XX 30-OCT-2002 (first entry)  
 XX  
 XX OmpA-K2S fusion protein encoding DNA.  
 DE  
 XX Tissue plasminogen activator; tPA; kringle 2 serine protease; stroke;  
 KW K2S; cardiac infarction; acute myocardial infarction; pulmonary embolism;  
 KW artery occlusion; deep vein thrombosis; blood clotting; thrombolytic;  
 KW cerebroprotective; cardiant; ompA; fusion protein; gene; ds.  
 XX  
 XX Unidentified.  
 OS  
 XX  
 XX Key Location/Qualifiers  
 FH 1..1128  
 CDS /\*tag= a  
 FT /product= "OmpA-K2S fusion protein"  
 FT /transl\_except= (pos:1126..1128, aa:Pro-Gly)  
 FT /note= "No stop codon"  
 FT /partial  
 XX  
 XX WO200240650-A2.  
 XX  
 XX 23-MAY-2002.  
 XX  
 XX 07-NOV-2001; 2001WO-EP12857.  
 XX  
 XX 14-NOV-2000; 2000GB-0027779.  
 XX  
 XX (BOEH ) BOEHRINGER INGELHEIM INT GMBH.  
 XX  
 XX Werner R, Goetz F, Tayapiwatana C, Manosroi J, Manosroi A;  
 XX WPI; 2002-471625/50.  
 XX

DR WPI; 2002-519376/55.  
 DR P-FSDB; AAE25034.  
 PT Producing active, correctly folded recombinant tissue plasminogen  
 PT activator, Kringle 2 serine protease in prokaryotic cells by expressing  
 PT the protein-encoding DNA operably linked to DNA coding for signal  
 PT peptide OmpA  
 XX  
 XX Claim 7; Page 30; 80pp; English.  
 PS  
 CC The present invention relates to a method of producing extracellularly  
 CC secreted, active, correctly folded, recombinant tissue plasminogen  
 CC activator (tPA), Kringle 2 serine protease molecule (K2S) or their  
 CC variants in prokaryotic cells by expressing the protein-encoding DNA  
 CC operably linked to DNA coding for signal peptide OmpA. The method is  
 CC useful for producing recombinant DNA-derived tPA, K2S or their variants.  
 CC Sequences of the invention are useful for manufacturing a medicament  
 CC for treating stroke, cardiac infarction, acute myocardial infarction,  
 CC pulmonary embolism, any artery occlusion such as intracranial artery  
 CC occlusion (e.g., arteries supplying the brain), peripherally occluded  
 CC arteries, coronary artery occlusion, deep vein thrombosis or related  
 CC diseases associated with unwanted blood clotting. The present sequence  
 CC is a DNA encoding a fusion protein comprising OmpA and K2S protein.  
 XX  
 XX Sequence 1128 BP; 238 A; 332 C; 332 G; 226 T; 0 other;  
 SQ

Query Match 100.0%; Score 66; DB 24; Length 1128;  
 Best Local Similarity 100.0%; Pred. No. 4.3e-14;  
 Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGAAAAGACAGCTATCGGATTGCGAGTGCAGTGCAGTGGCTGCTTTCGCTACCGTGGCCAG 60  
 Db 1 ATGAAAAGACAGCTATCGGATTGCGAGTGCAGTGCAGTGGCTGCTTTCGCTACCGTGGCCAG 60  
 Qy 61 GCGGCC 66  
 Db 61 GCGGCC 66

RESULT 4  
 AAZ57599  
 ID AAZ57599 standard; DNA; 1574 BP.  
 AC AAZ57599;  
 XX  
 XX 19-APR-2000 (first entry)  
 DT  
 XX Erythrocyte and Hepatitis B dual-specific antibody DNA sequence.  
 DE  
 XX Erythrocyte; Hepatitis B; dual-specific antibody; surface antigen; ds.  
 KW  
 XX Unidentified.  
 OS  
 XX CN1232039-A.  
 PN  
 XX 20-OCT-1999.  
 PD  
 XX 02-APR-1999; 99CN-0103517.  
 PF  
 XX 02-APR-1999; 99CN-0103517.  
 PR  
 XX (NAVA-) NAVAL GEN HOSPITAL PLA.  
 PA  
 XX Chen Y, Wang Y;  
 PI  
 XX WPI; 2000-098467/09.  
 DR  
 XX Genetic engineering double specific antibody and its use -  
 PT  
 XX Claim 1; Page 1-2; 6pp; Chinese.  
 PS  
 XX The present invention describes a dual-specific antibody composed of  
 CC the surface antigens to resist against erythrocyte and hepatitis B,

CC which can be used as the test reagent to detect the surface antigen  
 CC of hepatitis B in blood. The dual-specific antibody is prepared by  
 CC recombination technique in gene engineering. The recombination technique  
 CC includes reforming the surface antigen to resist against erythrocyte  
 CC and hepatitis B to become hybrid antibody genes by shortening the  
 CC joining peptide of single-chain antibody, and assembling them in the  
 CC same expression vector. The gene product can be directly extracted from  
 CC the supernatant of bacterial culture liquid. Its advantages are low  
 CC cost, quick detection and simple operation.  
 XX  
 XX Sequence 1574 BP; 395 A; 390 C; 419 G; 370 T; 0 other;  
 SQ

Query Match 100.0%; Score 66; DB 21; Length 1574;  
 Best Local Similarity 100.0%; Pred. No. 4.6e-14;  
 Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGAAAAGACAGCTATCGGATTGCGAGTGCAGTGCAGTGGCTGCTTTCGCTACCGTGGCCAG 60  
 Db 792 ATGAAAAGACAGCTATCGGATTGCGAGTGCAGTGCAGTGGCTGCTTTCGCTACCGTGGCCAG 851  
 Qy 61 GCGGCC 66  
 Db 852 GCGGCC 857

RESULT 5  
 AAC66993  
 ID AAC66993 standard; DNA; 3450 BP.  
 XX  
 XX AAC66993;  
 AC  
 XX 28-MAR-2001 (first entry)  
 DT  
 XX Vector pCGMT-1b.  
 DB  
 XX Filamentous phage; protein display; pVII; PIX;  
 KW combinatorial antibody library; ds.  
 XX  
 XX Synthetic.  
 OS  
 XX WO200071694-A1.  
 PN  
 XX 30-NOV-2000.  
 PD  
 XX 24-MAY-2000; 2000WO-US14433.  
 PF  
 XX 25-MAY-1999; 99US-0318786.  
 PR  
 XX (SCRI ) SCRIPPS RES INST.  
 XX  
 XX Janda KD, Wirsching P, Lerner RA, Gao C;  
 PI  
 XX WPI; 2001-032030/04.  
 DR  
 XX Novel filamentous phage encapsulating a genome encoding fusion  
 XX polypeptide comprising exogenous polypeptide fused to amino terminus of  
 PT pVII and PIX proteins, for constructing diverse heterodimeric  
 PT polypeptide array -  
 XX  
 XX Example 1; Page 74-75; 90pp; English.  
 PS  
 XX The present invention describes a filamentous phage encapsulating a  
 CC genome encoding a fusion protein. This fusion protein comprises an  
 CC exogenous protein fused to the amino terminus of a filamentous phage pVII  
 CC or PIX protein. This is useful in the design of proteins for medical,  
 CC industrial, environmental and research applications.  
 XX  
 XX Sequence 3450 BP; 873 A; 799 C; 891 G; 887 T; 0 other;  
 SQ

Query Match 100.0%; Score 66; DB 22; Length 3450;  
 Best Local Similarity 100.0%; Pred. No. 5.6e-14;  
 Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAAAAGACAGCTATCGGATTCAGTGGCACTGGCTGGTTTCGTACCGTGGCCCGAG 60  
 Db 2667 ATGAAAAGACAGCTATCGGATTCAGTGGCACTGGCTGGTTTCGTACCGTGGCCCGAG 2726

QY 61 GCGGCC 66

Db 2727 GCGGCC 2732

# RESULT 6

AAQ92546

ID AAQ92546 standard; DNA; 4691 BP.

XX

AC AAQ92546;

XX

DT 11-MAR-1996 (first entry)

XX

DE pComb3 expression vector.

XX

KW pComb3; phagemid expression vector; bacteriophage; coat protein 3;  
 KW Gene III; filamentous phage; minor phage coat protein; cpIII; cp3;  
 KW bacterial membrane; periplasm; E. coli; human; Fab; HIV; gp120;  
 KW combinatorial Fab library; cassette; Fd/cp3; lacZ promoter/operator;  
 KW ribosome binding site; RBS; PelB leader; spacer; tether sequence;  
 KW MT4; pMT4-3; antibody; ss; cyclic.

XX

OS Synthetic.

XX

PN WO9511317-A1.

XX

XX 27-APR-1995.

PD

XX 19-OCT-1994; 94WO-US11907.

PF

XX 19-SEP-1994; 94US-0308841.

PR

XX 19-OCT-1993; 93US-0139409.

PR

XX 26-APR-1994; 94US-0233619.

XX

PA (SCRI) SCRIPPS RES INST.

XX

PI Barbas CF, Burton DR, Lerner RA;

XX

XX WPI; 1995-170235/22.

XX

PT Synthetic human neutralising monoclonal antibodies to human

PT

PT immunodeficiency virus - used for diagnosis and immuno:therapy of

PT

XX HIV-induced disease

XX

PS Example 1; Page 185-188; 249pp; English.

XX

CC This sequence represents the pComb3 phagemid expression vector. This  
 CC vector has been designed to allow for anchoring of expressed proteins on  
 CC the bacteriophage coat protein 3. Gene III of filamentous phage encodes  
 CC the 406 residue minor phage coat protein, cpIII (cp3), which is  
 CC expressed prior to extrusion in the phage assembly process on a bacterial  
 CC membrane and accumulates on the inner membrane facing into the periplasm  
 CC of E. coli. This plasmid was used within the scope of the invention to  
 CC express various mutagenised human Fab's which comprise heavy and light  
 CC variable regions which bind to HIV gp120. pComb3 allows for both surface  
 CC cloning of soluble forms of the Fabs. The vector was designed for the  
 CC cloning of combinatorial Fab libraries. pComb consists of a DNA molecule  
 CC having two cassettes to express one fusion protein, Fd/cp3, and one  
 CC soluble protein, the light chain. The finished vector comprises,  
 CC operatively linked 5' to 3', a first cassette consisting of lacZ  
 CC promoter/operator sequences, a NotI restriction site, a ribosome binding  
 CC site (RBS), a PelB leader, a spacer region, a cloning region bordered by  
 CC 5' XhoI and 3' SpeI restriction sites, the tether sequence, the sequences  
 CC encoding bacteriophage cp3 followed by a stop codon, a XbaI restriction  
 CC site between the two cassettes, and a second lacZ promoter/operator  
 CC sequence, followed by an expression control RBS, a XbaI leader, a spacer  
 CC region, a cloning region bordered by 5' SacI and 3' XbaI restriction  
 CC sites, followed by expression control stop sequences and a second NotI  
 CC restriction site. The pComb3 expression vector forms the basic

CC construct of the MT4 Fab display phagemid expression vector, pMT4-3 (see  
 CC also AAQ92540), used in the invention for the production of synthetic  
 CC human Fab antibodies against gp120 of HIV.

SQ Sequence 4691 BP; 1170 A; 1171 C; 1232 G; 1118 T; 0 other;

Query Match 100.0%; Score 66; DB 16; Length 4691;  
 Best Local Similarity 100.0%; Pred. No. 6.1e-14;  
 Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAAAAGACAGCTATCGGATTCAGTGGCACTGGCTGGTTTCGTACCGTGGCCCGAG 60  
 Db 2611 ATGAAAAGACAGCTATCGGATTCAGTGGCACTGGCTGGTTTCGTACCGTGGCCCGAG 2670

QY 61 GCGGCC 66

Db 2671 GCGGCC 2676

# RESULT 7

ABQ73372

ID ABQ73372 standard; DNA; 5149 BP.

XX

AC ABQ73372;

XX

DT 01-OCT-2002 (first entry)

XX

DE Plasmid pRL8 nucleotide sequence SEQ ID NO:60.

XX

KW TPO; EPO; thrombopoietin; erythropoietin; antibody; CDR region;  
 KW complementarity determining region; immunoglobulin; antianaemic;  
 KW haemostatic; nephrotropic; haematopoietic cell; haematopoiesis;  
 KW gene; ds.

XX

OS Homo sapiens.

OS

XX Synthetic.

XX

PN WO200246238-A2.

XX

PD 13-JUN-2002.

XX

XX 05-DEC-2001; 2001WO-US47656.

XX

XX 05-DEC-2000; 2000US-251448P.

PR

XX 04-MAY-2001; 2001US-288889P.

PR

XX 29-MAY-2001; 2001US-294068P.

XX

PA (ALEX-) ALEXION PHARM INC.

XX

XX Bowdish KS, Barbas-Frederickson S, Renshaw M;

XX

XX WPI; 2002-566610/60.

XX

XX A novel immunogen molecule comprising a region in which amino acid

XX

XX residues corresponding to at least a portion of the complementary

XX

XX determining region are replaced or fused with an erythropoietin or  
 XX thrombopoietin mimetic -  
 XX

XX

PS Disclosure; Fig 6A-C; 113pp; English.

XX

XX The present invention describes an immunoglobulin molecule or its fragment  
 CC (I) comprising a region where amino acid residues corresponding to at  
 CC least a portion of the complementary determining region (CDR) are  
 CC replaced or fused with biologically active peptides e.g. a peptide  
 CC mimetic such as an erythropoietin (EPO) or thrombopoietin (TPO) mimetic,  
 CC that is flanked with proline at its carboxy terminus. (I) has  
 CC antianaemic, haemostatic and nephrotropic activities, and can be used as  
 CC a stimulator of proliferation, differentiation and maturation of  
 CC haematopoietic cells, and a stimulator of haematopoiesis. (I) is useful  
 CC for stimulating proliferation, differentiation or growth of  
 CC promegakaryocytes or megakaryocytes, where (I) is contacted with  
 CC promegakaryocytes or megakaryocytes, which results in increased platelet  
 CC production. (I) with a region where amino acid residues corresponding to

CC a portion of CDR is replaced with an EPO mimetic, or which has one or  
 CC more of its CDRs fused to an EPO mimetic, is useful for increasing the  
 CC production of red blood cells, where (I) is contacted with haematopoietic  
 CC stem cells or their progenitors. (I) is useful for diagnostics or  
 CC therapeutics, in cell isolation strategies, and for treating patients  
 CC suffering from deficiency in cell populations caused by disease,  
 CC disorders or treatments related to the suppression of haematopoiesis.  
 CC ABQ73288 to ABQ73377 and ABP51669 to ABP51696 represent sequences used  
 CC in the exemplification of the present invention.

SQ Sequence 5149 BP; 1279 A; 1244 C; 1372 G; 1254 T; 0 other;

Query Match 100.0%; Score 66; DB 24; Length 5149;  
 Best Local Similarity 100.0%; Pred. No. 6.2e-14;  
 Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGAAAAGACAGCTATCGCGATTGCGAGTGGCTGGCTGGCTTACCGTGGCCAG 60  
 Db 2611 ATGAAAAGACAGCTATCGCGATTGCGAGTGGCTGGCTTACCGTGGCCAG 2670

Qy 61 GCGGCC 66  
 Db 2671 GCGGCC 2676

RESULT 8  
 ABN84077  
 ID ABN84077 standard; DNA; 5149 BP.

AC ABN84077;

DT 23-SEP-2002 (first entry)

DE Phagemid vector pRL8.

KW Vector; phagemid; pRL8; pComb; dimerisation domain; Jun;  
 KW leucine zipper; antibody; gene; db.

OS Unidentified.

PH Key Location/Qualifiers

FT misc\_feature 2677..3866

FT /note= "light chain stuffer region"

FT /tag= a

FT misc\_feature 3968..4337

FT /tag= b

FT /note= "heavy chain stuffer region"

FT 4338..4583

FT /tag= c

FT /partial

FT /product= "Jun dimerisation domain"

FT /note= "the CDS does not include a start codon"

XX WO200246434-A2.

XX 13-JUN-2002.

XX 07-DEC-2001; 2001WO-US46514.

XX 08-DEC-2000; 2000US-254410P.

XX (ALEX-) ALEXION PHARM INC.

XX Bowdish KS, Barbas-Frederickson S;

XX WPI; 2002-537568/57.

XX P-PSDB; ABB79463.

XX Novel pComb family phagemid comprising a nucleic acid encoding a  
 PT dimerization domain, useful for producing antibody fragments which  
 PT require dimerization in order to crosslink for activation of target  
 PT receptors

XX

PS Disclosure; Fig 8A-C; 17pp; English.

XX The present sequence is of plasmid pRL8, a pComb vector family  
 CC member. pRL8 was constructed by insertion of a PCR fragment  
 CC containing a flexible linker (murine kappa hinge region) followed  
 CC by a Jun leucine zipper dimerisation domain into a pRL4 backbone  
 CC that had been treated to remove the stuffer region. The resulting  
 CC construct was then cut with restriction endonucleases to replace  
 CC scfv with a pRL4 stuffer region for the light and heavy chain. The  
 CC vector is useful in the production of biologically active proteins  
 CC such as antibody fragments which require dimerisation for their  
 CC biological activity. It can also be used in the production and  
 CC screening of libraries made in accordance with phage display  
 CC technology. Homo-dimerisation of single chain antibodies is  
 CC accomplished by the presence of the Jun leucine zipper domain,  
 CC which is responsible for protein-protein interactions. Subcloning  
 CC to pRL8 individually or en masse following FACS sorting or panning  
 CC allows expression, e.g., of dimeric soluble Fabs.

SQ Sequence 5149 BP; 1279 A; 1244 C; 1372 G; 1254 T; 0 other;

Query Match 100.0%; Score 66; DB 24; Length 5149;  
 Best Local Similarity 100.0%; Pred. No. 6.2e-14;  
 Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGAAAAGACAGCTATCGCGATTGCGAGTGGCTGGCTTACCGTGGCCAG 60  
 Db 2611 ATGAAAAGACAGCTATCGCGATTGCGAGTGGCTGGCTTACCGTGGCCAG 2670

Qy 61 GCGGCC 66

Db 2671 GCGGCC 2676

RESULT 9

AAL41116

ID AAL41116 standard; DNA; 5683 BP.

XX AAL41116;

DT 16-OCT-2002 (first entry)

XX Plasmid pRL5 nucleic acid sequence.

DE Plasmid; vector; replication; expression; foreign genetic; bacteria;  
 KW cyanobacterium; Escherichia coli; pComb3X plasmid; pRL5; pRL5 asc-CAT;  
 KW pRL5-CAT; pRL5 bsi-CAT; gene; db.

OS Unidentified.

PH Key Location/Qualifiers

FT CDS 591..1451

FT /tag= a

FT /product= "Beta lactamase"

FT 2611..2676

FT /tag= b

FT /product= "Omp A leader amino acid sequence"

FT 3508..3828

FT /tag= c

FT /product= "Kappa constant region amino acid sequence"

FT 3859..3924

FT /tag= d

FT /product= "pel B leader amino acid sequence"

FT 4740..5651

FT /tag= e

FT /product= "CH1, His6 tag, HA tag, and gene III amino acid  
 sequence"

XX WO200246436-A2.

XX 13-JUN-2002.

XX 07-DEC-2001; 2001WO-US46516.

XX

```

XX 08-DEC-2000; 2000US-254411P.
XX (ALEX-) ALEXION PHARM INC.
XX Bowdish KS, Barbas-frederickson S, Wild M, McWhirter J;
XX WPI; 2002-537570/57.
XX P-PSDB; AAO22536, AAO22537, AAO22538, AAO22539, AAO22542.
XX Novel plasmid useful in cloning and expression of foreign genetic
XX information -
XX Claim 1; Fig 4A-E; 39pp; English.
XX The invention relates to a plasmid with a nucleic acid sequence of 6122,
XX 6125, or 5683 base pairs fully defined in the specification. The
XX invention more specifically relates to novel vectors capable of
XX replication and expression of foreign genetic information in bacteria,
XX such as, for example, cyanobacterium and E. coli. The new vectors have
XX been designed to overcome certain drawbacks of the pComb3X plasmid. These
XX new vectors include pRL5, pRL5-CAT, pRL5 asc-CAT and pRL5 bsi-CAT. This
XX polynucleotide sequence represents the plasmid pRL5 nucleic acid
XX sequence of the invention.
XX Sequence 5683 BP; 1337 A; 1495 C; 1540 G; 1311 T; 0 other;
XX
XX Query Match 100.0%; Score 66; DB 24; Length 5683;
XX Best Local Similarity 100.0%; Pred. No. 6.4e-14;
XX Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 ATGAAAAGACAGCTATCGGATTCGAGTGGCAGTGGCTGTTTCGTACCGTGCCCGAG 60
XX AAL41112
XX DB 2611 ATGAAAAGACAGCTATCGGATTCGAGTGGCAGTGGCTGTTTCGTACCGTGCCCGAG 2670
XX
XX QY 61 GCGGCC 66
XX DB 2671 GCGGCC 2676
XX
XX RESULT 10
XX AAL41112
XX ID AAL41112 standard; DNA; 6122 BP.
XX AC AAL41112;
XX DT 16-OCT-2002 (first entry)
XX DE Plasmid pRL5 CAT nucleic acid sequence.
XX KW Plasmid; vector; replication; expression; foreign genetic; bacteria;
XX KW cyanobacterium; Escherichia coli; pComb3X plasmid; pRL5; pRL5 asc-CAT;
XX KW pRL5-CAT; pRL5 bsi-CAT; gene; ds.
XX OS Unidentified.
XX FH Location/Qualifiers
XX FT 1013..1672
XX FT /*tag= a
XX FT /product= "CAT amino acid sequence"
XX FT CDS
XX FT 3050..3115
XX FT /*tag= b
XX FT /product= "Omp A leader amino acid sequence"
XX FT CDS
XX FT 3953..4267
XX FT /*tag= c
XX FT /product= "Kappa constant region amino acid sequence"
XX FT CDS
XX FT 4298..4363
XX FT /*tag= d
XX FT /product= "pel B leader amino acid sequence"
XX FT CDS
XX FT 5179..6090
XX FT /*tag= e
XX FT /product= "CH1, His6 tag, HA tag, and gene III amino acid
XX FT sequence"

```

```

XX WO200246436-A2.
XX 13-JUN-2002.
XX 07-DEC-2001; 2001WO-US46516.
XX 08-DEC-2000; 2000US-254411P.
XX (ALEX-) ALEXION PHARM INC.
XX Bowdish KS, Barbas-frederickson S, Wild M, McWhirter J;
XX WPI; 2002-537570/57.
XX P-PSDB; AAO22535, AAO22536, AAO22537, AAO22538, AAO22539.
XX Novel plasmid useful in cloning and expression of foreign genetic
XX information -
XX Claim 1; Fig 7A-E; 39pp; English.
XX The invention relates to a plasmid with a nucleic acid sequence of 6122,
XX 6125, or 5683 base pairs fully defined in the specification. The
XX invention more specifically relates to novel vectors capable of
XX replication and expression of foreign genetic information in bacteria,
XX such as, for example, cyanobacterium and E. coli. The new vectors have
XX been designed to overcome certain drawbacks of the pComb3X plasmid. These
XX new vectors include pRL5, pRL5-CAT, pRL5 asc-CAT and pRL5 bsi-CAT. This
XX polynucleotide sequence represents the plasmid pRL5-CAT nucleic acid
XX sequence of the invention.
XX Sequence 6122 BP; 1449 A; 1571 C; 1639 G; 1463 T; 0 other;
XX
XX Query Match 100.0%; Score 66; DB 24; Length 6122;
XX Best Local Similarity 100.0%; Pred. No. 6.5e-14;
XX Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 ATGAAAAGACAGCTATCGGATTCGAGTGGCAGTGGCTGTTTCGTACCGTGCCCGAG 60
XX DB 3050 ATGAAAAGACAGCTATCGGATTCGAGTGGCAGTGGCTGTTTCGTACCGTGCCCGAG 3109
XX
XX QY 61 GCGGCC 66
XX DB 3110 GCGGCC 3115
XX
XX RESULT 11
XX AAL41115
XX ID AAL41115 standard; DNA; 6122 BP.
XX AC AAL41115;
XX DT 16-OCT-2002 (first entry)
XX DE Plasmid pRL5 bsi-CAT nucleic acid sequence.
XX KW Plasmid; vector; replication; expression; foreign genetic; bacteria;
XX KW cyanobacterium; Escherichia coli; pComb3X plasmid; pRL5; pRL5 asc-CAT;
XX KW pRL5-CAT; pRL5 bsi-CAT; gene; ds.
XX OS Unidentified.
XX FH Location/Qualifiers
XX FT 1013..1672
XX FT /*tag= a
XX FT /product= "CAT amino acid sequence"
XX FT CDS
XX FT 3050..3115
XX FT /*tag= b
XX FT /product= "Omp A leader amino acid sequence"
XX FT CDS
XX FT 3947..4267
XX FT /*tag= c
XX FT /product= "Kappa constant region amino acid sequence"
XX FT CDS
XX FT 4298..4363

```



```

FT      /*tag= d
FT      /product= "pel B leader amino acid sequence"
FT      5179..6090
FT      /*tag= e
FT      /product= "CH1, His6 tag, HA tag, and gene III amino acid
FT      sequence"
XX
FN      WO200246436-A2.
XX
PD      13-JUN-2002.
XX
PF      07-DEC-2001; 2001WO-US46516.
XX
PR      08-DEC-2000; 2000US-254411P.
XX
PA      (ALEX-) ALEXION PHARM INC.
XX
PI      Bowdish KS, Barbas-Frederickson S, Wild M, McWhirter J;
XX
WPI; 2002-537570/57.
DR      P-PSDB; AAO22535, AAO22536, AAO22538, AAO22539, AAO22541.
XX
XX      Novel plasmid useful in cloning and expression of foreign genetic
PT      information
XX
PS      Disclosure; Fig 11A-E; 39pp; English.
XX
CC      The invention relates to a plasmid with a nucleic acid sequence of 6122,
CC      6125, or 5683 base pairs fully defined in the specification. The
CC      invention more specifically relates to novel vectors capable of
CC      replication and expression of foreign genetic information in bacteria,
CC      such as, for example, cyanobacterium and E. coli. The new vectors have
CC      been designed to overcome certain drawbacks of the pComb3X plasmid. These
CC      new vectors include pRL5, pRL5-CAT, pRL5 asc-CAT and pRL5 bsi-CAT. This
CC      polynucleotide sequence represents the plasmid pRL5 bsi-CAT nucleic acid
CC      sequence of the invention.
XX
SQ      Sequence 6122 BP; 1447 A; 1571 C; 1641 G; 1463 T; 0 other;

Query Match          100.0%; Score 66; DB 24; Length 6122;
Best Local Similarity 100.0%; Pred. No. 6.5e-14;
Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 ATGAAAAGACAGCTATCGGATTCAGTGCAGTGGCTGGTTTCGTACCGTGGCCGAC 60
Db      3050 ATGAAAAGACAGCTATCGGATTCAGTGCAGTGGCTGGTTTCGTACCGTGGCCGAC 3109
QY      61 GCGGCC 66
Db      3110 GCGGCC 3115

RESULT 12
ABN84078
ID      ABN84078 standard; DNA; 6122 BP.
XX
AC      ABN84078;
XX
DT      23-SEP-2002 (first entry)
XX
DE      Plasmid pRL5-CAT.
XX
KW      Plasmid pRL5-CAT; vector; antibody; gene; ds.
XX
OS      Unidentified.
XX
FH      Key
FT      CDS
FT      Location/Qualifiers
FT      591..800
FT      /*tag= a
FT      /partial
FT      /product= "Amp frag"
FT      /note= "the CDS does not include a stop codon"
FT      1013..1672
FT      CDS

```

```

FT      /*tag= b
FT      /product= "Chloramphenicol transferase"
FT      complement (2052..2651)
FT      /*tag= c
FT      /note= "ori."
FT      2831..3046
FT      /*tag= d
FT      /note= "lac promoter"
FT      3009..3029
FT      /*tag= e
FT      /note= "lac rep site"
FT      3036..3039
FT      /*tag= f
FT      3050..3115
FT      /*tag= g
FT      /partial
FT      /product= "OmpA leader"
FT      /note= "the CDS does not include a stop codon"
FT      3121..3946
FT      /*tag= h
FT      /note= "light chain variable region stuffer"
FT      3951..4269
FT      /*tag= i
FT      /note= "Kappa Cns"
FT      4298..4363
FT      /*tag= j
FT      /partial
FT      /product= "pelB leader"
FT      /note= "the CDS does not include a stop codon"
FT      4385..5147
FT      /*tag= k
FT      /note= "heavy chain variable region stuffer"
FT      5485..5556
FT      /*tag= l
FT      /partial
FT      /product= "Linker-His tag-HA tag"
FT      /note= "the CDS does not include a start codon"
FT      5557..6090
FT      /*tag= m
FT      /partial
FT      /gene= "Gene III"
FT      /note= "the CDS does not include a start codon"
XX
XX      WO200246435-A2.
XX      13-JUN-2002.
XX
XX      05-DEC-2001; 2001WO-US47452.
XX
XX      05-DEC-2000; 2000US-251440P.
XX      (ALEX-) ALEXION PHARM INC.
XX
XX      Bowdish KS, Barbas-Frederickson S, Lin Y, Renshaw M, Wild M;
XX      McWhirter J;
XX
XX      WPI; 2002-537569/57.
XX
XX      Plasmid for in situ production of genes, comprises two template
XX      annealing sequences, such as downstream primer and upstream collar
XX      sequence and a restriction site located between the annealing sequences
XX
XX      Disclosure; Fig 4A-T; 65pp; English.
XX
XX      The present sequence is that of plasmid pRL5-CAT, a derivative of
XX      pComb 3X which has been modified to contain chloramphenicol
XX      resistance. The invention relates to the engineering of plasmids,
XX      of which pRL-CAT is a preferred example, for in situ production of
XX      genes. It was found that nucleic acids encoding a polypeptide
XX      can be directly incorporated into a plasmid by DNA polymerisation
XX      or by reverse transcription of a nucleic acid template. Preferably,
XX      nucleic acids encoding at least a portion of an antibody can be

```

CC directly incorporated into the plasmid by reverse transcription of  
 CC mRNA. The plasmids are engineered to contain 2 template annealing  
 CC sequences (see ABN84079-80 and ABN84082-83), i.e. a downstream  
 CC primer that anneals to a first portion of a nucleic acid template,  
 CC e.g. mRNA encoding at least a portion of an antibody, an upstream  
 CC collar sequence that anneals to a second portion of the template,  
 CC and at least 1 restriction site located between the 2 template  
 CC annealing sequences. A single-stranded DNA plasmid vector is  
 CC produced containing a nucleic acid encoding at least a portion of  
 CC a polypeptide, e.g. a light chain and/or a heavy chain of an  
 CC antibody. This vector can be transformed into a host cell and  
 CC amplified.

XX Sequence 6122 BP; 1449 A; 1571 C; 1639 G; 1463 T; 0 other;  
 SQ Query Match 100.0%; Score 66; DB 24; Length 6122;  
 Best Local Similarity 100.0%; Pred. No. 6.5e-14;  
 Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAAAAGACAGCTATCGCGATTGCGAGTGGCACTGGCTGGTTTCGCTACCGTGCCCGAG 60  
 DB 3050 ATGAAAAGACAGCTATCGCGATTGCGAGTGGCACTGGCTGGTTTCGCTACCGTGCCCGAG 3109  
 QY 61 GCGGCC 66  
 DB 3110 GCGGCC 3115

RESULT 13  
 AAL41114  
 ID AAL41114 standard; DNA; 6125 BP.  
 AC AAL41114;  
 XX  
 DT 16-OCT-2002 (first entry)  
 DE Plasmid pRL5 CAT-Asc nucleic acid sequence.  
 KW Plasmid; vector; replication; expression; foreign genetic; bacteria;  
 KW cyanobacterium; Escherichia coli; pComb3X plasmid; pRL5; pRL5 asc-CAT;  
 KW pRL5-CAT; pRL5 bsi-CAT; gene; ds.  
 XX  
 OS Unidentified.  
 FH Key Location/Qualifiers  
 FT CDS 1013..1672  
 FT /\*tag= a  
 FT /product= "CAT amino acid sequence"  
 FT 3050..3115  
 FT /\*tag= b  
 FT /product= "Omp A leader amino acid sequence"  
 FT 3953..4270  
 FT /\*tag= c  
 FT /product= "Kappa constant region amino acid sequence"  
 FT 4301..4366  
 FT /\*tag= d  
 FT /product= "pel B leader amino acid sequence"  
 FT 5182..6093  
 FT /\*tag= e  
 FT /product= "CHI1, His6 tag, HA tag, and gene III amino acid  
 FT sequence"  
 XX WO200246436-A2.  
 PN  
 XX  
 PD 13-JUN-2002.  
 XX  
 XX 07-DEC-2001; 2001WO-US46516.  
 XX PF  
 XX 08-DEC-2000; 2000US-254411P.  
 XX PR  
 XX (ALEX-) ALEXION PHARM INC.  
 XX PA  
 XX Bowdish KS, Barbas-frederickson S, Wild M, McWhirter J;

XX WPI; 2002-537570/57.  
 DR P-PSDB; AAO22535, AAO22536, AAO22538, AAO22539, AAO22540.  
 XX  
 FT Novel plasmid useful in cloning and expression of foreign genetic  
 FT information  
 XX  
 PS Claim 1; Fig 9A-E; 39pp; English.  
 XX  
 CC The invention relates to a plasmid with a nucleic acid sequence of 6122,  
 CC 6125, or 5683 base pairs fully defined in the specification. The  
 CC invention more specifically relates to novel vectors capable of  
 CC replication and expression of foreign genetic information in bacteria,  
 CC such as, for example, cyanobacterium and E. coli. The new vectors have  
 CC been designed to overcome certain drawbacks of the pComb3X plasmid. These  
 CC new vectors include pRL5, pRL5-CAT, pRL5 asc-CAT and pRL5 bsi-CAT. This  
 CC polynucleotide sequence represents the plasmid pRL5 CAT-Asc nucleic acid  
 CC sequence of the invention.  
 XX  
 SQ Sequence 6125 BP; 1447 A; 1573 C; 1644 G; 1461 T; 0 other;  
 Query Match 100.0%; Score 66; DB 24; Length 6125;  
 Best Local Similarity 100.0%; Pred. No. 6.5e-14;  
 Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAAAAGACAGCTATCGCGATTGCGAGTGGCACTGGCTGGTTTCGCTACCGTGCCCGAG 60  
 DB 3050 ATGAAAAGACAGCTATCGCGATTGCGAGTGGCACTGGCTGGTTTCGCTACCGTGCCCGAG 3109  
 QY 61 GCGGCC 66  
 DB 3110 GCGGCC 3115

RESULT 14  
 AAO92547  
 ID AAO92547 standard; DNA; 6166 BP.  
 AC AAO92547;  
 XX  
 DT 11-MAR-1996 (first entry)  
 DE Expression vector, pPho-TT.  
 XX  
 KW Human; Fab; variable chain; heavy; light; region; VH; VL; HIV; gp120;  
 KW 3b1; 3b3; 3b4; 3b9; WT4; humanised; monoclonal antibody; MAb;  
 KW immunoreaction; neutralisation; passive immunotherapy; tetanus toxin;  
 KW alkaline phosphatase; phoA; ss; cyclic.  
 XX  
 OS Synthetic.  
 XX WO9511317-A1.  
 PN  
 XX  
 PD 27-APR-1995.  
 XX  
 XX 19-OCT-1994; 94WO-US11907.  
 PF  
 XX 19-SEP-1994; 94US-0308841.  
 PR  
 XX 19-OCT-1993; 93US-0139409.  
 PR  
 XX 26-APR-1994; 94US-0233619.  
 PR  
 XX (SCRI ) SCRIPPS RES INST.  
 PA  
 XX  
 XX Barbas CF, Burton DR, Lerner RA;  
 PI  
 XX WPI; 1995-170235/22.  
 XX  
 XX Synthetic human neutralising monoclonal antibodies to human  
 XX immunodeficiency virus - used for diagnosis and immuno:therapy of  
 XX HIV-induced disease  
 XX  
 XX Example 2; Page 193-197; 249pp; English.  
 XX

CC This Omp A signal peptide-encoding sequence is contained in a



Result No.	Score	Query Match	Length	DB	ID	Description	§
1	66	100.0	3450	4	US-09-318-786-20	Sequence 20, Appl	
2	66	100.0	4691	3	US-08-591-633-43	Sequence 43, Appl	
3	66	100.0	4691	4	US-09-611-451-43	Sequence 43, Appl	
4	66	100.0	6166	3	US-08-591-632-51	Sequence 51, Appl	
5	66	100.0	6166	4	US-09-611-451-51	Sequence 51, Appl	
6	60.2	91.2	73	1	US-07-854-845B-2	Sequence 2, Appl	
7	60.2	91.2	77	3	US-08-666-354A-10	Sequence 10, Appl	
8	60.2	91.2	249	1	US-08-439-132-1	Sequence 1, Appl	
9	60.2	91.2	249	6	5223407-1	Patent No. 5223407	
10	60.2	91.2	548	3	US-08-828-741B-1	Sequence 1, Appl	
11	60.2	91.2	548	4	US-09-160-567-1	Sequence 1, Appl	
12	60.2	91.2	548	4	US-09-710-299-1	Sequence 1, Appl	
13	59.6	90.3	458	2	US-08-880-829-21	Sequence 21, Appl	
14	59.6	90.3	470	3	US-08-828-741B-10	Sequence 10, Appl	
15	59.6	90.3	470	4	US-09-160-567-10	Sequence 10, Appl	
16	59.6	90.3	470	4	US-09-710-299-10	Sequence 10, Appl	
17	59.6	90.3	599	3	US-08-828-741B-7	Sequence 7, Appl	
18	59.6	90.3	599	4	US-09-160-567-7	Sequence 7, Appl	
19	59.6	90.3	599	4	US-09-710-299-7	Sequence 7, Appl	
20	59.6	90.3	1031	3	US-08-828-741B-5	Sequence 5, Appl	
21	59.6	90.3	1031	4	US-09-160-567-5	Sequence 5, Appl	
22	59.6	90.3	1031	4	US-09-710-299-5	Sequence 5, Appl	
23	59.6	90.3	1490	3	US-08-828-741B-3	Sequence 3, Appl	
24	59.6	90.3	1490	4	US-09-160-567-3	Sequence 3, Appl	
25	59.6	90.3	1490	4	US-09-710-299-3	Sequence 3, Appl	
26	59.2	89.7	116	1	US-07-842-089E-23	Sequence 23, Appl	
27	59.2	89.7	116	1	US-08-264-485-23	Sequence 23, Appl	

; CORRESPONDENCE ADDRESS

ADDRESSEE: The Scripps Research Institute, Office of  
ADDRESSER: Patent Counsel  
STREET: 10550 No. 6261558th Torrey Pines Road, TPC 8  
CITY: La Jolla  
STATE: CA  
COUNTRY: USA  
ZIP: 92037

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/591,632  
FILING DATE: 19-OCT-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US94/11907  
FILING DATE: 19-OCT-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/308,841  
FILING DATE: 19-SEP-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/233,619  
FILING DATE: 26-APR-1994

APPLICATION NUMBER: US 08/139,409  
FILING DATE: 19-OCT-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Fitting, Thomas  
REGISTRATION NUMBER: 34,163  
REFERENCE/DOCKET NUMBER: TSRI 332.3  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619) 784-2937  
TELEFAX: (619) 784-9399

INFORMATION FOR SEQ ID NO: 43:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4691 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: circular  
MOLECULE TYPE: DNA (genomic)  
US-08-591-632-43

Query Match 100.0%; Score 66; DB 3; Length 4691;  
Best Local Similarity 100.0%; Pred. No. 1.le-13;  
Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGAAAAGACAGCTATCGGATTGCGAGTGCAGTGGCTGTTTGGCTACCGTGGCCCGAG 60  
Db 2611 ATGAAAAGACAGCTATCGGATTGCGAGTGCAGTGGCTGTTTGGCTACCGTGGCCCGAG 2670

Qy 61 GCGGCC 66  
Db 2671 GCGGCC 2676

## RESULT 3

US-09-611-451-43  
Sequence 43, Application US/09611451  
Patent No. 6395275  
GENERAL INFORMATION:  
APPLICANT: Barbas, Carlos F.  
Burton, Dennis R.  
Lerner, Richard A.

TITLE OF INVENTION: SYNTHETIC HUMAN NEUTRALIZING MONOCLONAL  
ANTIBODIES TO HUMAN IMMUNODEFICIENCY VIRUS

NUMBER OF SEQUENCES: 92  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: The Scripps Research Institute, Office of  
Patent Counsel

STREET: 10550 No. 6395275th Torrey Pines Road, TPC 8  
CITY: La Jolla  
STATE: CA

COUNTRY: USA  
ZIP: 92037  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/611,451  
FILING DATE: 06-Jul-2000

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/591,632  
FILING DATE: 2001-10-29  
APPLICATION NUMBER: US 08/308,841  
FILING DATE: 19-SEP-1994  
APPLICATION NUMBER: US 08/233,619  
FILING DATE: 26-APR-1994  
APPLICATION NUMBER: US 08/139,409  
FILING DATE: 19-OCT-1993

ATTORNEY/AGENT INFORMATION:  
NAME: Fitting, Thomas  
REGISTRATION NUMBER: 34,163  
REFERENCE/DOCKET NUMBER: TSRI 332.3  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619) 784-2937  
TELEFAX: (619) 784-9399

INFORMATION FOR SEQ ID NO: 43:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4691 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: circular

MOLECULE TYPE: DNA (genomic)  
SEQUENCE DESCRIPTION: SEQ ID NO: 43:  
US-09-611-451-43

Query Match 100.0%; Score 66; DB 4; Length 4691;  
Best Local Similarity 100.0%; Pred. No. 1.le-13;  
Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGAAAAGACAGCTATCGGATTGCGAGTGCAGTGGCTGTTTGGCTACCGTGGCCCGAG 60  
Db 2611 ATGAAAAGACAGCTATCGGATTGCGAGTGCAGTGGCTGTTTGGCTACCGTGGCCCGAG 2670

Qy 61 GCGGCC 66  
Db 2671 GCGGCC 2676

## RESULT 4

US-08-591-632-51  
Sequence 51, Application US/08591632  
Patent No. 6261558  
GENERAL INFORMATION:  
APPLICANT: Barbas, Carlos F.  
Burton, Dennis R.

APPLICANT: Lerner, Richard A.  
TITLE OF INVENTION: SYNTHETIC HUMAN NEUTRALIZING MONOCLONAL  
ANTIBODIES TO HUMAN IMMUNODEFICIENCY VIRUS

NUMBER OF SEQUENCES: 92  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: The Scripps Research Institute, Office of  
Patent Counsel

STREET: 10550 No. 6261558th Torrey Pines Road, TPC 8  
CITY: La Jolla  
STATE: CA  
COUNTRY: USA  
ZIP: 92037

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0

CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US 08/591,632  
 FILING DATE: 19-OCT-1994  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: PCT/US94/11907  
 FILING DATE: 19-OCT-1994  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/308,841  
 FILING DATE: 19-SEP-1994  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/233,619  
 FILING DATE: 26-APR-1994  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/139,409  
 FILING DATE: 19-OCT-1993  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Fitting, Thomas  
 REGISTRATION NUMBER: 34,163  
 REFERENCE/DOCKET NUMBER: TSRI 332.3  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (619) 784-2937  
 TELEFAX: (619) 784-9399  
 INFORMATION FOR SEQ ID NO: 51:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 6166 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: circular  
 MOLECULE TYPE: DNA (genomic)  
 US-08-591-632-51

Query Match 100.0%; Score 66; DB 3; Length 6166;  
 Best Local Similarity 100.0%; Pred. No. 1.2e-13;  
 Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAAAACAGCAGCTATCGGATTCGAGTTCGAGTGCAGTGGCTGTTTCGCTACCGTGCCCGAG 60  
 DB 4541 ATGAAAACAGCAGCTATCGGATTCGAGTTCGAGTGCAGTGGCTGTTTCGCTACCGTGCCCGAG 4600

QY 61 GCGGCC 66  
 DB 4601 GCGGCC 4606

RESULT 5  
 US-09-611-451-51  
 Sequence 51, Application US/09611451  
 Patent No. 6395275  
 GENERAL INFORMATION:  
 APPLICANT: Barbas, Carlos F.  
 Burton, Dennis R.  
 Lerner, Richard A.  
 TITLE OF INVENTION: SYNTHETIC HUMAN NEUTRALIZING MONOCLONAL ANTIBODIES TO HUMAN IMMUNODEFICIENCY VIRUS  
 NUMBER OF SEQUENCES: 92  
 CORRESPONDENCE ADDRESS:  
 ADDRESSES: The Scripps Research Institute, Office of Patent Counsel  
 STREET: 10550 No. 6395275th Torrey Pines Road, TPC 8  
 CITY: La Jolla  
 STATE: CA  
 COUNTRY: USA  
 ZIP: 92037  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent In Release #1.0  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/611,451  
 FILING DATE: 06-Jul-2000  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 08/591,632

FILING DATE: 2001-10-29  
 APPLICATION NUMBER: US 08/308,841  
 FILING DATE: 19-SEP-1994  
 APPLICATION NUMBER: US 08/233,619  
 FILING DATE: 26-APR-1994  
 APPLICATION NUMBER: US 08/139,409  
 FILING DATE: 19-OCT-1993  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Fitting, Thomas  
 REGISTRATION NUMBER: 34,163  
 REFERENCE/DOCKET NUMBER: TSRI 332.3  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (619) 784-2937  
 TELEFAX: (619) 784-9399  
 INFORMATION FOR SEQ ID NO: 51:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 6166 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: circular  
 MOLECULE TYPE: DNA (genomic)  
 SEQUENCE DESCRIPTION: SEQ ID NO: 51:  
 US-09-611-451-51

Query Match 100.0%; Score 66; DB 4; Length 6166;  
 Best Local Similarity 100.0%; Pred. No. 1.2e-13;  
 Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAAAACAGCAGCTATCGGATTCGAGTTCGAGTGCAGTGGCTGTTTCGCTACCGTGCCCGAG 60  
 DB 4541 ATGAAAACAGCAGCTATCGGATTCGAGTTCGAGTGCAGTGGCTGTTTCGCTACCGTGCCCGAG 4600

QY 61 GCGGCC 66  
 DB 4601 GCGGCC 4606

RESULT 6  
 US-07-854-845B-2  
 Sequence 2, Application US/07854845B  
 Patent No. 5340732  
 GENERAL INFORMATION:  
 APPLICANT: HABUKA, No. 5340732iyuki  
 APPLICANT: MIYANO, Masashi  
 APPLICANT: MATSUMOTO, Takashi  
 APPLICANT: NOMA, Masana  
 TITLE OF INVENTION: ANTIVIRAL PROTEIN  
 NUMBER OF SEQUENCES: 9  
 CORRESPONDENCE ADDRESS:  
 ADDRESSES: BIRCH, STEWART, KOLASCH, & BIRCH  
 STREET: 8110 Gatehouse Road, Suite 500 East  
 CITY: Falls Church  
 STATE: VA  
 COUNTRY: USA  
 ZIP: 22042  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent In Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/07/854,845B  
 FILING DATE: 20-MAR-1992  
 CLASSIFICATION: 530  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Murphy Jr., Gerald M  
 REGISTRATION NUMBER: 28,977  
 REFERENCE/DOCKET NUMBER: 42-194P  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 703-205-8000  
 TELEFAX: 703-205-8050  
 TELEX: 248345  
 INFORMATION FOR SEQ ID NO: 2:

## SEQUENCE CHARACTERISTICS:

LENGTH: 73 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
US-07-854-845B-2

Query Match 91.2%; Score 60.2; DB 1; Length 73;  
Best Local Similarity 95.4%; Pred. No. 4.9e-12;  
Matches 62; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ATGAAAAGACAGCTATCGGATTCAGTGGCACTGGCTGTTTCGCTACCGTGCCCGCAG 60

Db 2 ATGAAAAGACAGCTATCGGATTCAGTGGCACTGGCTGTTTCGCTACCGTGCCCGCAG 61

QY 61 GCGGC 65

Db 62 GCGGC 66

## RESULT 7

US-08-666-354A-10  
Sequence 10, Application US/08666354A  
Patent No. 6040141

## GENERAL INFORMATION:

APPLICANT: KLAUSER, THOMAS  
APPLICANT: KRAMER, JOACHIM  
APPLICANT: MEYER, THOMAS F.  
APPLICANT: POHLNER, JOHANNES

TITLE OF INVENTION: BACTERIA USED TO PRODUCE STABLE FUSION

TITLE OF INVENTION: PROTEINS AND METHOD FOR THEIR IDENTIFICATION

NUMBER OF SEQUENCES: 18

CORRESPONDENCE ADDRESS:

ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH, LLP.

STREET: P.O. BOX 747

CITY: FALLS CHURCH

STATE: VA

COUNTRY: USA

ZIP: 22040-0747

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/666,354A

FILING DATE: 23-SEP-1996

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: SVENSSON, LEONARD R.

REGISTRATION NUMBER: 30,330

REFERENCE/DOCKET NUMBER: 147-157P

TELECOMMUNICATION INFORMATION:

TELEPHONE: 703-205-8000

TELEFAX: 703-205-8050

INFORMATION FOR SEQ ID NO: 10:

SEQUENCE CHARACTERISTICS:

LENGTH: 77 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: other nucleic acid

DESCRIPTION: /desc = "FRAGMENT OF PLASMID pJK165"

US-08-666-354A-10

Query Match 91.2%; Score 60.2; DB 3; Length 77;

Best Local Similarity 95.4%; Pred. No. 4.9e-12;

Matches 62; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ATGAAAAGACAGCTATCGGATTCAGTGGCACTGGCTGTTTCGCTACCGTGCCCGCAG 60

Db 1 ATGAAAAGACAGCTATCGGATTCAGTGGCACTGGCTGTTTCGCTACCGTAGCGCAG 60

QY 61 GCGGC 65

Db 61 GCGGC 65

## RESULT 8

US-08-439-132-1

Sequence 1, Application US/08439132

Patent No. 5646015

GENERAL INFORMATION:

APPLICANT: Wong, W. K. R.

APPLICANT: Sutherland, Margaret L.

TITLE OF INVENTION: EXCRETION OF HETEROLOGOUS PROTEINS FROM

TITLE OF INVENTION: E. COLI

NUMBER OF SEQUENCES: 5

CORRESPONDENCE ADDRESS:

ADDRESSEE: Foley & Lardner

STREET: Suite 500, 3000 K Street

CITY: Washington

STATE: D.C.

COUNTRY: USA

ZIP: 20007-5109

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/439,132

FILING DATE: 11-MAY-1995

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Bent, Stephen A.

REGISTRATION NUMBER: 29,768

REFERENCE/DOCKET NUMBER: 16777/200/ALLE

TELECOMMUNICATION INFORMATION:

TELEPHONE: (202) 672-5300

TELEFAX: (202) 672-5399

TELEX: 904136

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 249 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

US-08-439-132-1

Query Match 91.2%; Score 60.2; DB 1; Length 249;

Best Local Similarity 95.4%; Pred. No. 6.2e-12;

Matches 62; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ATGAAAAGACAGCTATCGGATTCAGTGGCACTGGCTGTTTCGCTACCGTGCCCGCAG 60

Db 102 ATGAAAAGACAGCTATCGGATTCAGTGGCACTGGCTGTTTCGCTACCGTAGCGCAG 161

QY 61 GCGGC 65

Db 162 GCGGC 166

## RESULT 9

5223407-1

Patent No. 5223407

APPLICANT: WONG, RAYMOND W.K.; SUTHERLAND, MARGARET L.

TITLE OF INVENTION: EXCRETION OF HETEROLOGOUS PROTEINS

FROM E. COLI

NUMBER OF SEQUENCES: 6

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/395,797

FILING DATE: 18-AUG-1989





US-09-710-299-1  
; Sequence 1, Application US/09710299  
; Patent No. 6521741  
; GENERAL INFORMATION:  
; APPLICANT: Koentgen, Frank  
; Suesse, Gabriele M.  
; Tarlinton, David M.  
; Treutlein, Herbert R.  
; TITLE OF INVENTION: CATALYTIC ANTIBODIES AND A METHOD OF  
; PRODUCING SAME  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER  
; STREET: 400 Garden City Plaza  
; CITY: Garden City  
; STATE: New York  
; COUNTRY: United States of America  
; ZIP: 11530  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/710,299  
; FILING DATE: 09-No. 6521741-2000  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/828,741  
; FILING DATE: <Unknown>  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Digiglio, Frank S.  
; REGISTRATION NUMBER: 31,346  
; REFERENCE/DOCKET NUMBER: 10591  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (516) 742-4343  
; TELEFAX: (516) 742-4366  
; TELEX: 230 901 SANS UR  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 548 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 1..548  
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
US-09-710-299-1

Query Match 91.2%; Score 60.2; DB 4; Length 548;  
Best Local Similarity 95.4%; Pred. No. 7.2e-12;  
Matches 62; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ATGAAAAGACAGCTATCGGATTCAGTGGCAGTGGCTGTTTCGCTACCGTGCAG 60  
Db 1 ATGAAAAGACAGCTATCGGATTCAGTGGCAGTGGCTGTTTCGCTACCGTGCAG 60

Qy 61 GCGGC 65

Db 61 GCGGC 65

RESULT 13  
US-08-880-829-21  
; Sequence 21, Application US/08880829  
; Patent No. 5925559  
; GENERAL INFORMATION:  
; APPLICANT: Collins, John  
; APPLICANT: Roettgen, Peter  
; TITLE OF INVENTION: A Collection of Phagemids, A  
; TITLE OF INVENTION: Collection of Escherichia Coli

; TITLE OF INVENTION: Cells Carrying The Phagemids, A  
; TITLE OF INVENTION: Collection of Phagemid Particles  
; TITLE OF INVENTION: Produced From Said Collection  
; TITLE OF INVENTION: And Phagemid Particles  
; NUMBER OF SEQUENCES: 21  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Joseph T. Eisele  
; ADDRESSEE: Kane, Dalsimer, Sullivan, Kurucz,  
; ADDRESSEE: Levy, Eisele and Richard  
; STREET: 711 Third Avenue  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10017-4059  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3-1/2" DISKETTE  
; COMPUTER: IBM-XT COMPATIBLE  
; OPERATING SYSTEM: DOS 3.3;  
; SOFTWARE: WORDPERFECT 5.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/880,829  
; FILING DATE: 23-JUN-1997  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/458,668  
; FILING DATE: 06/02/95  
; APPLICATION NUMBER: German EP 94 108 689.4  
; FILING DATE: 06/07/94  
; ATTORNEY/AGENT INFORMATION:  
; NAME: EISELE, JOSEPH T.  
; REGISTRATION NUMBER: 25,331  
; REFERENCE/DOCKET NUMBER: 2727-77  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 687-6000  
; TELEFAX: (212) 682-3485  
; TELEX: (212) 426767  
; INFORMATION FOR SEQ ID NO: 21:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 458 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single strand  
; TOPOLOGY: linear  
; MOLECULE TYPE:  
; HYPOTHETICAL:  
; ANTI-SENSE:  
; FRAGMENT TYPE:  
; ORIGINAL SOURCE:  
; ORGANISM:  
; STRAIN:  
; INDIVIDUAL ISOLATE:  
; DEVELOPMENTAL STAGE:  
; HAPLOTYPE:  
; TISSUE TYPE:  
; CELL TYPE:  
; CELL LINE:  
; ORGANELLE:  
; IMMEDIATE SOURCE:  
US-08-880-829-21

Query Match 90.3%; Score 59.6; DB 2; Length 458;  
Best Local Similarity 93.9%; Pred. No. 1.1e-11;  
Matches 62; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 ATGAAAAGACAGCTATCGGATTCAGTGGCAGTGGCTGTTTCGCTACCGTGCAG 60  
Db 112 ATGAAAAGACAGCTATCGGATTCAGTGGCAGTGGCTGTTTCGCTACCGTGCAG 171

Qy 61 GCGGCC 66

Db 172 GCGGAC 177

RESULT 14

US-08-828-741B-10  
; Sequence 10, Application US/08828741B  
; Patent No. 6043069  
; GENERAL INFORMATION:  
; APPLICANT: Koentgen, Frank  
; APPLICANT: Sues, Gabriele M.  
; APPLICANT: Tarlington, David M.  
; APPLICANT: Treutlein, Herbert R.  
; TITLE OF INVENTION: CATALYTIC ANTIBODIES AND A METHOD OF  
; TITLE OF INVENTION: PRODUCING SAME  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER  
; STREET: 400 Garden City Plaza  
; CITY: Garden City  
; STATE: New York  
; COUNTRY: United States of America  
; ZIP: 11530  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/828,741B  
; FILING DATE: 26-MAR-1997  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Digiglio, Frank S.  
; REGISTRATION NUMBER: 31,346  
; REFERENCE/DOCKET NUMBER: 10591  
; TELEPHONE: (516) 742-4343  
; TELEFAX: (516) 742-4366  
; TELEX: 230 901 SANS UR  
; INFORMATION FOR SEQ ID NO: 10:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 470 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 1..470  
US-08-828-741B-10  
Query Match 90.3%; Score 59.6; DB 3; Length 470;  
Best Local Similarity 93.9%; Pred. No. 1.le-11;  
Matches 62; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
Qy 1 ATGAAAAGACAGCTATCGGATTCGACGTGGCTGGTTCGCTACCGTGCCCGAG 60  
Db 1 ATGAAAAGACAGCTATCGGATTCGACGTGGCTGGTTCGCTACCGTGCCCGAG 60  
Qy 61 GCGGCC 66  
Db 61 GCCGAC 66  
RESULT 15  
US-09-160-567-10  
; Sequence 10, Application US/09160567  
; Patent No. 6326179  
; GENERAL INFORMATION:  
; APPLICANT: Koentgen, Frank  
; APPLICANT: Sues, Gabriele M.  
; APPLICANT: Tarlington, David M.  
; APPLICANT: Treutlein, Herbert R.  
; TITLE OF INVENTION: CATALYTIC ANTIBODIES AND A METHOD OF  
; TITLE OF INVENTION: PRODUCING SAME  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:

; ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER  
; STREET: 400 Garden City Plaza  
; CITY: Garden City  
; STATE: New York  
; COUNTRY: United States of America  
; ZIP: 11530  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/160,567  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/828,741  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Digiglio, Frank S.  
; REGISTRATION NUMBER: 31,346  
; REFERENCE/DOCKET NUMBER: 10591  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (516) 742-4343  
; TELEFAX: (516) 742-4366  
; TELEX: 230 901 SANS UR  
; INFORMATION FOR SEQ ID NO: 10:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 470 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 1..470  
US-09-160-567-10

Query Match 90.3%; Score 59.6; DB 4; Length 470;  
Best Local Similarity 93.9%; Pred. No. 1.le-11;  
Matches 62; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
Qy 1 ATGAAAAGACAGCTATCGGATTCGACGTGGCTGGTTCGCTACCGTGCCCGAG 60  
Db 1 ATGAAAAGACAGCTATCGGATTCGACGTGGCTGGTTCGCTACCGTGCCCGAG 60  
Qy 61 GCGGCC 66  
Db 61 GCCGAC 66

Search completed: August 19, 2003, 08:35:08  
Job time : 7.16534 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 19, 2003, 04:46:44 ; Search time 18.6243 Seconds  
(without alignments)  
7930.701 Million cell updates/sec

Title: US-09-987-455-3

Perfect score: 66

Sequence: 1 atgaaaagacagctatcgc.....ctaccgtggccagcgccg 66

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1504479 seqs, 1118970152 residues

Total number of hits satisfying chosen parameters: 3008958

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:\*

- 1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/ptodata/2/pubpna/ECT\_NEW\_PUB.seq:\*
- 3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:\*
- 4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:\*
- 5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:\*
- 6: /cgn2\_6/ptodata/2/pubpna/ECTUS\_PUBCOMB.seq:\*
- 7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:\*
- 8: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq:\*
- 9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq:\*
- 10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq:\*
- 11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq:\*
- 12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:\*
- 13: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq:\*
- 14: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq:\*
- 15: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:\*
- 16: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:\*
- 17: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	66	100.0	66	11	US-09-987-457-1
2	66	100.0	66	11	US-09-987-455-3
3	66	100.0	66	11	US-09-987-455-6
4	66	100.0	1128	11	US-09-987-455-2
5	66	100.0	1128	11	US-09-987-455-5
6	66	100.0	5149	14	US-10-006-593-60
7	66	100.0	6122	14	US-10-006-591-1
8	60.2	91.2	548	12	US-10-345-618-1
9	59.6	90.3	470	12	US-10-345-618-10
10	59.6	90.3	599	12	US-10-345-618-7
11	59.6	90.3	932	9	US-09-809-517A-39
12	59.6	90.3	1031	12	US-10-345-618-5
13	59.6	90.3	1479	12	US-10-345-618-15
14	59.6	90.3	1490	12	US-10-345-618-3
15	59.6	90.3	1574	9	US-09-809-517A-38
16	59.6	90.3	4425	9	US-09-809-517A-40

17	59.2	89.7	4145	14	US-10-001-934-36	Sequence 36, Appl
18	59.2	89.7	5020	14	US-10-001-934-35	Sequence 35, Appl
19	59.2	89.7	5079	9	US-09-809-517A-41	Sequence 41, Appl
20	58.8	89.1	63	9	US-09-760-008A-4	Sequence 4, Appl
21	58.8	89.1	63	14	US-10-076-117-3	Sequence 3, Appl
22	58.8	89.1	63	14	US-10-192-294-4	Sequence 4, Appl
23	58.8	89.1	65	9	US-09-875-494-20	Sequence 20, Appl
24	58.8	89.1	108	11	US-09-848-616-8	Sequence 8, Appl
25	58.8	89.1	256	10	US-09-916-230-13	Sequence 13, Appl
26	58.8	89.1	256	11	US-09-848-616-18	Sequence 18, Appl
27	58.8	89.1	261	10	US-09-916-230-15	Sequence 15, Appl
28	58.8	89.1	261	11	US-09-848-616-20	Sequence 20, Appl
29	58.8	89.1	4614	9	US-09-912-165-17	Sequence 17, Appl
30	58.8	89.1	4657	9	US-09-912-165-18	Sequence 18, Appl
31	58	87.9	585	10	US-09-848-585-28	Sequence 28, Appl
32	58	87.9	585	10	US-09-848-585-30	Sequence 30, Appl
33	58	87.9	981	9	US-09-334-477-36	Sequence 36, Appl
34	58	87.9	990	9	US-09-334-477-38	Sequence 38, Appl
35	58	87.9	2321	10	US-09-995-396-2	Sequence 2, Appl
36	58	87.9	2337	10	US-09-995-396-3	Sequence 3, Appl
37	57.6	87.3	7083	10	US-09-995-396-1	Sequence 1, Appl
38	57.6	87.3	819	12	US-09-226-157-3	Sequence 3, Appl
39	57.2	86.7	903	14	US-10-033-399B-11	Sequence 11, Appl
40	55.4	83.9	93	13	US-10-004-832-3	Sequence 3, Appl
41	52.8	80.0	867	11	US-09-782-397-16	Sequence 16, Appl
c 42	52.8	80.0	867	11	US-09-782-397-18	Sequence 18, Appl
c 43	52.8	80.0	918	11	US-09-782-397-13	Sequence 13, Appl
c 44	52.8	80.0	918	11	US-09-782-397-15	Sequence 15, Appl
45	50	75.8	102	14	US-10-006-593-107	Sequence 107, App

## ALIGNMENTS

## RESULT 1

US-09-987-457-1  
; Sequence 1, Application US/09987457  
; Publication No. US20030013150A1  
; GENERAL INFORMATION:  
; APPLICANT: Manosroi, Aranya  
; APPLICANT: Manosroi, Aranya  
; APPLICANT: Tayapiwatana, Chatchai  
; APPLICANT: Goetz, Friedrich  
; APPLICANT: Werner, Rolf-Guenther  
; TITLE OF INVENTION: Methods for Large Scale Protein Production in Prokaryotes  
; FILE REFERENCE: 0652.2180001  
; CURRENT APPLICATION NUMBER: US/09/987,457  
; CURRENT FILING DATE: 2001-11-14  
; PRIOR APPLICATION NUMBER: 60/288,573  
; PRIOR FILING DATE: 2001-02-15  
; PRIOR APPLICATION NUMBER: GB 00 27 782.2  
; PRIOR FILING DATE: 2000-11-14  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 66  
; TYPE: DNA  
; ORGANISM: Escherichia coli  
US-09-987-457-1

Query Match 100.0%; Score 66; DB 11; Length 66;  
Best Local Similarity 100.0%; Pred. No. 4.6e-16;  
Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	ATGAAAAGACAGCTATCGCGATTGCAGTTCGCTGGCTGGCTTCGTTTCGTTACCGTGGCCAG	60
Db	1	ATGAAAAGACAGCTATCGCGATTGCAGTTCGCTGGCTGGCTTCGTTTCGTTACCGTGGCCAG	60
Qy	61	CGCGCC	66
Db	61	CGCGCC	66

RESULT 2  
US-09-987-455-3  
; Sequence 3, Application US/09987455  
; Publication No. US20030049729A1  
; GENERAL INFORMATION:  
; APPLICANT: Aranya Manosroi  
; APPLICANT: Jiradej Manosroi  
; APPLICANT: Chatchai Tayapiwatana  
; APPLICANT: Friedrich Goetz  
; APPLICANT: Rolf-Guenther Werner  
; TITLE OF INVENTION: Methods for Large Scale Production of Recombinant  
; FILE REFERENCE: 0652.2190001  
; CURRENT APPLICATION NUMBER: US/09/987,455  
; CURRENT FILING DATE: 2001-11-14  
; PRIOR APPLICATION NUMBER: 60/268,574  
; PRIOR FILING DATE: 2001-02-15  
; PRIOR APPLICATION NUMBER: GB 0027779.8  
; PRIOR FILING DATE: 2000-11-14  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 3  
; LENGTH: 66  
; TYPE: DNA  
; ORGANISM: Escherichia coli  
US-09-987-455-3

Query Match 100.0%; Score 66; DB 11; Length 66;  
Best Local Similarity 100.0%; Pred. No. 4.6e-16;  
Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGAAAAGACAGCTATCGCGATTGCGACTGGCTGGTTTCGCTACCGTGCCCGAG 60  
Db 1 ATGAAAAGACAGCTATCGCGATTGCGACTGGCTGGTTTCGCTACCGTGCCCGAG 60

Qy 61 GCGGCC 66  
Db 61 GCGGCC 66

RESULT 3  
US-09-987-455-6  
; Sequence 6, Application US/09987455  
; Publication No. US20030049729A1  
; GENERAL INFORMATION:  
; APPLICANT: Aranya Manosroi  
; APPLICANT: Jiradej Manosroi  
; APPLICANT: Chatchai Tayapiwatana  
; APPLICANT: Friedrich Goetz  
; APPLICANT: Rolf-Guenther Werner  
; TITLE OF INVENTION: Methods for Large Scale Production of Recombinant  
; FILE REFERENCE: 0652.2190001  
; CURRENT APPLICATION NUMBER: US/09/987,455  
; CURRENT FILING DATE: 2001-11-14  
; PRIOR APPLICATION NUMBER: 60/268,574  
; PRIOR FILING DATE: 2001-02-15  
; PRIOR APPLICATION NUMBER: GB 0027779.8  
; PRIOR FILING DATE: 2000-11-14  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 6  
; LENGTH: 66  
; TYPE: DNA  
; ORGANISM: Escherichia coli  
US-09-987-455-6

Query Match 100.0%; Score 66; DB 11; Length 66;  
Best Local Similarity 100.0%; Pred. No. 4.6e-16;  
Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGAAAAGACAGCTATCGCGATTGCGACTGGCTGGTTTCGCTACCGTGCCCGAG 60  
Db 1 ATGAAAAGACAGCTATCGCGATTGCGACTGGCTGGTTTCGCTACCGTGCCCGAG 60

Qy 61 GCGGCC 66  
Db 61 GCGGCC 66

Db 1 ATGAAAAGACAGCTATCGCGATTGCGACTGGCTGGTTTCGCTACCGTGCCCGAG 60  
Qy 61 GCGGCC 66  
Db 61 GCGGCC 66

RESULT 4  
US-09-987-455-2  
; Sequence 2, Application US/09987455  
; Publication No. US20030049729A1  
; GENERAL INFORMATION:  
; APPLICANT: Aranya Manosroi  
; APPLICANT: Jiradej Manosroi  
; APPLICANT: Chatchai Tayapiwatana  
; APPLICANT: Friedrich Goetz  
; APPLICANT: Rolf-Guenther Werner  
; TITLE OF INVENTION: Methods for Large Scale Production of Recombinant  
; FILE REFERENCE: 0652.2190001  
; CURRENT APPLICATION NUMBER: US/09/987,455  
; CURRENT FILING DATE: 2001-11-14  
; PRIOR APPLICATION NUMBER: 60/268,574  
; PRIOR FILING DATE: 2001-02-15  
; PRIOR APPLICATION NUMBER: GB 0027779.8  
; PRIOR FILING DATE: 2000-11-14  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 1128  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: coding  
; OTHER INFORMATION: sequence for OmpA-K2S fusion protein  
US-09-987-455-2

Query Match 100.0%; Score 66; DB 11; Length 1128;  
Best Local Similarity 100.0%; Pred. No. 9.7e-16;  
Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGAAAAGACAGCTATCGCGATTGCGACTGGCTGGTTTCGCTACCGTGCCCGAG 60  
Db 1 ATGAAAAGACAGCTATCGCGATTGCGACTGGCTGGTTTCGCTACCGTGCCCGAG 60

Qy 61 GCGGCC 66  
Db 61 GCGGCC 66

RESULT 5  
US-09-987-455-5  
; Sequence 5, Application US/09987455  
; Publication No. US20030049729A1  
; GENERAL INFORMATION:  
; APPLICANT: Aranya Manosroi  
; APPLICANT: Jiradej Manosroi  
; APPLICANT: Chatchai Tayapiwatana  
; APPLICANT: Friedrich Goetz  
; APPLICANT: Rolf-Guenther Werner  
; TITLE OF INVENTION: Methods for Large Scale Production of Recombinant  
; FILE REFERENCE: 0652.2190001  
; CURRENT APPLICATION NUMBER: US/09/987,455  
; CURRENT FILING DATE: 2001-11-14  
; PRIOR APPLICATION NUMBER: 60/268,574  
; PRIOR FILING DATE: 2001-02-15  
; PRIOR APPLICATION NUMBER: GB 0027779.8  
; PRIOR FILING DATE: 2000-11-14  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 5  
; LENGTH: 1128

```
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: coding
; OTHER INFORMATION: sequence for OmpA-K2S fusion protein
US-09-987-455-5

Query Match      100.0%; Score 66; DB 11; Length 1128;
Best Local Similarity 100.0%; Pred. No. 9.7e-16;
Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAAAAGACAGCTATCGGATTCAGTGGCACTGGCTGGTTTCGCTACCGTGCCCCAG 60
Db 1 ATGAAAAGACAGCTATCGGATTCAGTGGCACTGGCTGGTTTCGCTACCGTGCCCCAG 60

QY 61 GCGGCC 66
Db 61 GCGGCC 66

RESULT 6
US-10-006-593-60
; Sequence 60, Application US/10006593
; Publication No. US20030049683A1
; GENERAL INFORMATION:
; APPLICANT: Bowdish, Katherine S.
; APPLICANT: Frederickson, Shana
; APPLICANT: Renshaw, Mark
; TITLE OF INVENTION: RATIONALLY DESIGNED ANTIBODIES
; FILE REFERENCE: 1087-2
; CURRENT APPLICATION NUMBER: US/10/006,593
; CURRENT FILING DATE: 2001-12-05
; PRIOR APPLICATION NUMBER: US 60/251,448
; PRIOR FILING DATE: 2000-12-05
; PRIOR APPLICATION NUMBER: US 60/288,889
; PRIOR FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/294,068
; PRIOR FILING DATE: 2001-05-29
; NUMBER OF SEQ ID NOS: 118
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 60
; LENGTH: 5149
; TYPE: DNA
; ORGANISM: artificial sequence
; FEATURE:
; OTHER INFORMATION: vector
US-10-006-593-60

Query Match      100.0%; Score 66; DB 14; Length 5149;
Best Local Similarity 100.0%; Pred. No. 1.5e-15;
Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAAAAGACAGCTATCGGATTCAGTGGCACTGGCTGGTTTCGCTACCGTGCCCCAG 60
Db 2611 ATGAAAAGACAGCTATCGGATTCAGTGGCACTGGCTGGTTTCGCTACCGTGCCCCAG 2670

QY 61 GCGGCC 66
Db 2671 GCGGCC 2676

RESULT 7
US-10-006-591-1
; Sequence 1, Application US/10006591
; Publication No. US20030049731A1
; GENERAL INFORMATION:
; APPLICANT: Bowdish, Katherine S.
; APPLICANT: Frederickson, Shana
; APPLICANT: Lin, Ying-Chi
; APPLICANT: Renshaw, Mark
; APPLICANT: Wild, Martha
; APPLICANT: McWhirter, John
; TITLE OF INVENTION: ENGINEERED PLASMIDS AND THEIR USE FOR IN SITU PRODUCTION OF GENES

; FILE REFERENCE: 1087-3
; CURRENT APPLICATION NUMBER: US/10/006,591
; CURRENT FILING DATE: 2001-12-05
; PRIOR APPLICATION NUMBER: 60/251,440
; PRIOR FILING DATE: 2000-12-05
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 6122
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: vector
US-10-006-591-1

Query Match      100.0%; Score 66; DB 14; Length 6122;
Best Local Similarity 100.0%; Pred. No. 1.5e-15;
Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAAAAGACAGCTATCGGATTCAGTGGCACTGGCTGGTTTCGCTACCGTGCCCCAG 60
Db 3050 ATGAAAAGACAGCTATCGGATTCAGTGGCACTGGCTGGTTTCGCTACCGTGCCCCAG 3109

QY 61 GCGGCC 66
Db 3110 GCGGCC 3115

RESULT 8
US-10-345-618-1
; Sequence 1, Application US/10345618
; Publication No. US20030148484A1
; GENERAL INFORMATION:
; APPLICANT: Koentgen, Frank
; APPLICANT: Suess, Gabriele M.
; APPLICANT: Tarlincon, David M.
; APPLICANT: Treutlein, Herbert R.
; TITLE OF INVENTION: CATALYTIC ANTIBIOTIDES AND A METHOD OF PRODUCING SAME
; FILE REFERENCE: 13474
; CURRENT APPLICATION NUMBER: US/10/345,618
; CURRENT FILING DATE: 2003-01-16
; PRIOR APPLICATION NUMBER: US/09/509,031
; PRIOR FILING DATE: 2000-06-09
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 548
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(546)
; OTHER INFORMATION: Description of Artificial Sequence: LHL nucleotide
; OTHER INFORMATION: sequence
US-10-345-618-1

Query Match      91.2%; Score 60.2; DB 12; Length 548;
Best Local Similarity 95.4%; Pred. No. 1.4e-13;
Matches 62; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ATGAAAAGACAGCTATCGGATTCAGTGGCACTGGCTGGTTTCGCTACCGTGCCCCAG 60
Db 1 ATGAAAAGACAGCTATCGGATTCAGTGGCACTGGCTGGTTTCGCTACCGTGCCCCAG 60

QY 61 GCGGC 65
Db 61 GCGGC 65

RESULT 9
US-10-345-618-10
; Sequence 10, Application US/10345618
; Publication No. US20030148484A1
```

```
; GENERAL INFORMATION:
; APPLICANT: Koentgen, Frank
; APPLICANT: Sues, Gabriele M.
; APPLICANT: Tarlington, David M.
; APPLICANT: Treutlein, Herbert R.
; TITLE OF INVENTION: CATALYTIC ANTIBODIES AND A METHOD OF PRODUCING SAME
; FILE REFERENCE: 13474
; CURRENT APPLICATION NUMBER: US/10/345,618
; CURRENT FILING DATE: 2003-01-16
; PRIOR APPLICATION NUMBER: US/09/509,031
; PRIOR FILING DATE: 2000-06-09
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 470
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(468)
; OTHER INFORMATION: Description of Artificial Sequence:Kappa
; OTHER INFORMATION: nucleotide sequence
US-10-345-618-10

Query Match          90.3%; Score 59.6; DB 12; Length 470;
Best Local Similarity 93.9%; Pred. No. 2.3e-13;
Matches 62; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 ATGAAAAGACAGCTATCGGATTGCGAGTCGAGTGGCACTGGCTGGTTTCGCTACCGTGCCCGCAG 60
Db 1 ATGAAAAGACAGCTATCGGATTGCGAGTCGAGTGGCACTGGCTGGTTTCGCTACCGTGCCCGCAG 60

Qy 61 GCGGCC 66
Db 61 GCCGAC 66

RESULT 10
US-10-345-618-7
; Sequence 7, Application US/10345618
; Publication No. US20030148484A1
; GENERAL INFORMATION:
; APPLICANT: Koentgen, Frank
; APPLICANT: Sues, Gabriele M.
; APPLICANT: Tarlington, David M.
; APPLICANT: Treutlein, Herbert R.
; TITLE OF INVENTION: CATALYTIC ANTIBODIES AND A METHOD OF PRODUCING SAME
; FILE REFERENCE: 13474
; CURRENT APPLICATION NUMBER: US/10/345,618
; CURRENT FILING DATE: 2003-01-16
; PRIOR APPLICATION NUMBER: US/09/509,031
; PRIOR FILING DATE: 2000-06-09
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 599
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(597)
; OTHER INFORMATION: Description of Artificial Sequence:LHL.seq
; OTHER INFORMATION: nucleotide sequence
US-10-345-618-7

Query Match          90.3%; Score 59.6; DB 12; Length 599;
Best Local Similarity 93.9%; Pred. No. 2.5e-13;
Matches 62; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 ATGAAAAGACAGCTATCGGATTGCGAGTCGAGTGGCACTGGCTGGTTTCGCTACCGTGCCCGCAG 60
Db 1 ATGAAAAGACAGCTATCGGATTGCGAGTCGAGTGGCACTGGCTGGTTTCGCTACCGTGCCCGCAG 60
```

```
Qy 61 GCGGCC 66
Db 61 GCCGAC 66

RESULT 11
US-09-809-517A-39
; Sequence 39, Application US/09809517A
; Patent No. US20020034733A1
; GENERAL INFORMATION:
; APPLICANT: Lohning, Corinna
; TITLE OF INVENTION: No. US20020034733A1el methods for displaying (poly)peptides/prote
; TITLE OF INVENTION: particles via disulfide bonds
; FILE REFERENCE: MORPHO/11
; CURRENT APPLICATION NUMBER: US/09/809,517A
; CURRENT FILING DATE: 2001-03-15
; PRIOR APPLICATION NUMBER: EP 99114072.4
; PRIOR FILING DATE: 1999-07-20
; PRIOR APPLICATION NUMBER: EP 00103551.8
; PRIOR FILING DATE: 2000-02-18
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 39
; LENGTH: 932
; TYPE: DNA
; ORGANISM: artificial sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: expression cassette
US-09-809-517A-39

Query Match          90.3%; Score 59.6; DB 9; Length 932;
Best Local Similarity 93.9%; Pred. No. 2.8e-13;
Matches 62; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 ATGAAAAGACAGCTATCGGATTGCGAGTCGAGTGGCACTGGCTGGTTTCGCTACCGTGCCCGCAG 60
Db 266 ATGAAAAGACAGCTATCGGATTGCGAGTCGAGTGGCACTGGCTGGTTTCGCTACCGTGCCCGCAG 325

Qy 61 GCGGCC 66
Db 326 GCCGAC 331

RESULT 12
US-10-345-618-5
; Sequence 5, Application US/10345618
; Publication No. US20030148484A1
; GENERAL INFORMATION:
; APPLICANT: Koentgen, Frank
; APPLICANT: Sues, Gabriele M.
; APPLICANT: Tarlington, David M.
; APPLICANT: Treutlein, Herbert R.
; TITLE OF INVENTION: CATALYTIC ANTIBODIES AND A METHOD OF PRODUCING SAME
; FILE REFERENCE: 13474
; CURRENT APPLICATION NUMBER: US/10/345,618
; CURRENT FILING DATE: 2003-01-16
; PRIOR APPLICATION NUMBER: US/09/509,031
; PRIOR FILING DATE: 2000-06-09
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 1031
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1029)
; OTHER INFORMATION: Description of Artificial Sequence:TLHL nucleotide
; OTHER INFORMATION: sequence
US-10-345-618-5

Query Match          90.3%; Score 59.6; DB 12; Length 1031;
Best Local Similarity 93.9%; Pred. No. 2.9e-13;
```



```
Matches 62; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
Qy 1 ATGAAAAGACAGCTATCGGATTGCGACTGCGCTGGCTGGTTTCGTACCGTGGCCCGAG 60
    |||||
Db 1 ATGAAAAGACAGCTATCGGATTGCGACTGCGCTGGCTGGTTTCGTACCGTAGCGAG 60
    |||||
Qy 61 GCGGCC 66
    |||||
Db 61 GCCGAC 66

RESULT 13
US-10-345-618-15
; Sequence 15, Application US/10345618
; Publication No. US20030148484A1
; GENERAL INFORMATION:
; APPLICANT: Koentgen, Frank
; APPLICANT: Suess, Gabriele M.
; APPLICANT: Tarlinton, David M.
; APPLICANT: Treutlein, Herbert R.
; TITLE OF INVENTION: CATALYTIC ANTIBODIES AND A METHOD OF PRODUCING SAME
; FILE REFERENCE: 13474
; CURRENT APPLICATION NUMBER: US/10/345,618
; PRIOR FILING DATE: 2003-01-16
; PRIOR APPLICATION NUMBER: US/09/509,031
; PRIOR FILING DATE: 2000-06-09
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 15
; LENGTH: 1479
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: ccMTIgL
; OTHER INFORMATION: nucleotide sequence
US-10-345-618-15

Query Match 90.3%; Score 59.6; DB 12; Length 1479;
Best Local Similarity 93.9%; Pred. No. 3.2e-13;
Matches 62; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
Qy 1 ATGAAAAGACAGCTATCGGATTGCGACTGCGCTGGCTGGTTTCGTACCGTGGCCCGAG 60
    |||||
Db 28 ATGAAAAGACAGCTATCGGATTGCGACTGCGCTGGCTGGTTTCGTACCGTAGCGAG 87
    |||||
Qy 61 GCGGCC 66
    |||||
Db 88 GCCGAC 93

RESULT 14
US-10-345-618-3
; Sequence 3, Application US/10345618
; Publication No. US20030148484A1
; GENERAL INFORMATION:
; APPLICANT: Koentgen, Frank
; APPLICANT: Suess, Gabriele M.
; APPLICANT: Tarlinton, David M.
; APPLICANT: Treutlein, Herbert R.
; TITLE OF INVENTION: CATALYTIC ANTIBODIES AND A METHOD OF PRODUCING SAME
; FILE REFERENCE: 13474
; CURRENT APPLICATION NUMBER: US/10/345,618
; CURRENT FILING DATE: 2003-01-16
; PRIOR APPLICATION NUMBER: US/09/509,031
; PRIOR FILING DATE: 2000-06-09
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 1490
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: CDS
```

```
; LOCATION: (1)..(1488)
; OTHER INFORMATION: Description of Artificial Sequence: CATAB-TEV
; OTHER INFORMATION: nucleotide sequence
US-10-345-618-3

Query Match 90.3%; Score 59.6; DB 12; Length 1490;
Best Local Similarity 93.9%; Pred. No. 3.2e-13;
Matches 62; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
Qy 1 ATGAAAAGACAGCTATCGGATTGCGACTGCGCTGGCTGGTTTCGTACCGTGGCCCGAG 60
    |||||
Db 1 ATGAAAAGACAGCTATCGGATTGCGACTGCGCTGGCTGGTTTCGTACCGTAGCGAG 60
    |||||
Qy 61 GCGGCC 66
    |||||
Db 61 GCCGAC 66

RESULT 15
US-09-809-517A-38
; Sequence 38, Application US/09809517A
; Patent No. US20020034733A1
; GENERAL INFORMATION:
; APPLICANT: Lohning, Corinna
; TITLE OF INVENTION: No. US20020034733A1 methods for displaying (poly)peptides/prote
; FILE REFERENCE: MORPHO/11
; CURRENT APPLICATION NUMBER: US/09/809,517A
; CURRENT FILING DATE: 2001-03-15
; PRIOR APPLICATION NUMBER: EP 99114072.4
; PRIOR FILING DATE: 1999-07-20
; PRIOR APPLICATION NUMBER: EP 00103551.8
; PRIOR FILING DATE: 2000-02-18
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 38
; LENGTH: 1574
; TYPE: DNA
; ORGANISM: artificial sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: expression cassette
US-09-809-517A-38

Query Match 90.3%; Score 59.6; DB 9; Length 1574;
Best Local Similarity 93.9%; Pred. No. 3.2e-13;
Matches 62; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
Qy 1 ATGAAAAGACAGCTATCGGATTGCGACTGCGCTGGCTGGTTTCGTACCGTGGCCCGAG 60
    |||||
Db 266 ATGAAAAGACAGCTATCGGATTGCGACTGCGCTGGCTGGTTTCGTACCGTAGCGAG 325
    |||||
Qy 61 GCGGCC 66
    |||||
Db 326 GCCGAC 331

Search completed: August 19, 2003, 14:22:23
Job time : 19.6243 secs
```



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 18, 2003, 22:44:33 ; Search time 193.082 Seconds  
(without alignments)  
8307.845 Million cell updates/sec

Title: US-09-987-455-3  
Perfect score: 66  
Sequence: 1 atgaaaagacagctatgc.....ctaccgtggccagcgccg 66

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues  
Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST:  
1: em\_estba:\*  
2: em\_estum:\*  
3: em\_estin:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_htc:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_htc:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: em\_gss\_hum:\*  
18: em\_gss\_inv:\*  
19: em\_gss\_pln:\*  
20: em\_gss\_vrt:\*  
21: em\_gss\_fun:\*  
22: em\_gss\_mam:\*  
23: em\_gss\_mus:\*  
24: em\_gss\_pro:\*  
25: em\_gss\_rod:\*  
26: em\_gss\_phg:\*  
27: em\_gss\_vrl:\*  
28: gb\_gss1:\*  
29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	48	72.7	655	28	AF075794
C 2	27.8	42.1	386	14	CB806580
C 3	26.8	40.6	310	9	AL635506
C 4	26.8	40.6	312	9	AW004057

C 5	26.8	40.6	321	10	BE503913
C 6	26.8	40.6	325	9	AA398159
C 7	26.8	40.6	340	9	AA401699
C 8	26.8	40.6	409	13	BY641012
C 9	26.8	40.6	569	13	BU783334
C 10	26.8	40.6	667	13	EX099539
C 11	26.8	40.6	694	29	CNS04PH1
C 12	26.8	40.6	786	12	BI560274
C 13	26.8	40.6	871	13	BUS53590
C 14	26.6	40.3	268	28	AQ050715
C 15	26.6	40.3	359	9	AU244728
C 16	26.6	40.3	377	9	AU244560
C 17	26.6	40.3	450	10	BG234261
C 18	26.6	40.3	472	28	AQ045335
C 19	26.6	40.3	612	12	BJ031326
C 20	26.6	40.3	645	13	BQ400448
C 21	26.6	40.3	753	14	CA808042
C 22	26.6	40.3	833	14	CB349272
C 23	26.6	40.3	885	13	BQ737041
C 24	26.6	40.3	888	13	BQ737333
C 25	26.6	40.3	889	14	CD101043
C 26	26.6	40.3	897	13	BQ733411
C 27	26.6	40.3	980	14	CB206009
C 28	26.6	40.3	1124	14	CB206058
C 29	26.2	39.7	271	9	AU184837
C 30	26.2	39.7	374	13	BQ094120
C 31	26.2	39.7	695	14	CD355744
C 32	26.2	39.7	973	13	BQ938965
C 33	26	39.4	495	28	AQ833969
C 34	26	39.4	853	29	CNS03TOR
C 35	26	39.4	950	29	CNS04IET
C 36	26	39.4	1016	29	CNS02VHS
C 37	26	39.4	1387	29	BZ557231
C 38	25.6	38.8	291	14	H33495
C 39	25.6	38.8	331	14	CA377775
C 40	25.6	38.8	438	14	CA373964
C 41	25.6	38.8	621	9	AL966295
C 42	25.6	38.8	652	9	AL644617
C 43	25.6	38.8	673	9	AL655729
C 44	25.6	38.8	687	14	CA369617
C 45	25.6	38.8	691	9	AL630991

ALIGNMENTS

RESULT 1  
AF075794/c  
LOCUS AF075794 Salmonella typhimurium LT2, Lambda DASH II Salmonella typhimurium genomic clone 107-T3, genomic survey sequence. 655 bp DNA linear GSS 29-AUG-2000  
DEFINITION AF075794 Salmonella typhimurium LT2, Lambda DASH II Salmonella typhimurium genomic clone 107-T3, genomic survey sequence.  
ACCESSION AF075794  
VERSION AF075794.1 GI:3320664  
KEYWORDS GSS.  
SOURCE Salmonella typhimurium  
ORGANISM Salmonella typhimurium  
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales; Enterobacteriaceae; Salmonella.  
REFERENCE 1 (bases 1 to 655)  
Wong,R.M.-Y., Wong,K.K., Benson,N.R. and McClelland,M.  
TITLE Sample sequencing of a Salmonella typhimurium LT2 lambda library: comparison to the Escherichia coli K12 genome  
JOURNAL FEMS Microbiol. Lett. 173 (2), 411-423 (1999)  
MEDLINE 9243757  
PubMed 10227170  
COMMENT Contact: McClelland M  
Molecular Biology  
Sidney Kimmel Cancer Center  
3099 Science Park Road, San Diego, CA 92121, USA  
Email: mclelland@lifsci.sdsu.edu  
Class: shotgun.  
Location/Qualifiers  
FEATURES  
source 1. .655

/organism="Salmonella typhimurium"  
/mol\_type="genomic DNA"  
/strain="LT2"  
/db\_xref="taxon:602"  
/clone="107-T3"  
/clone\_lib="Salmonella typhimurium LT2, Lambda DASH II"  
/note="Vector: Lambda DASH II; sequenced using Li-Cor  
sequencer"  
BASE COUNT 195 a 158 c 139 g 161 t 2 others  
ORIGIN  
Query Match 72.7%; Score 48; DB 28; Length 655;  
Best Local Similarity 85.2%; Pred. No. 8.9e-06;  
Matches 52; Conservative 1; Mismatches 8; Indels 0; Gaps 0;  
QY 5 AAAAGACAGCTATCGGATTCAGTGGCACTGGCTGGTTTCGCTACCGTGCCCGGCGG 64  
Db 298 AAAAGACAGCTATCGGATTCAGTGGCACTGGCTGGTTTCGCTACCGTGCCCGGCGG 239  
QY 65 C 65  
Db 238 C 238

RESULT 2  
CB806580/c  
LOCUS  
DEFINITION AMGNNUC:SRP92-00054-G9-A srp92 (10238) Rattus norvegicus cDNA clone  
srp92-00054-g9 5', mRNA sequence.

ACCESSION CB806580

VERSION CB806580

KEYWORDS EST.

SOURCE Rattus norvegicus (Norway rat)

ORGANISM  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.

REFERENCE 1 (bases 1 to 386)

AUthORS Amgen EST Program.

TITLE Amgen Rat EST Program

JOURNAL Unpublished

COMMENT Contact: Dan Fitzpatrick

One Amgen Center Drive, Thousand Oaks, CA 91320-1799, USA

Tel: 805 447-4881

Plate: 00054 row; g column: 9.

Location/Qualifiers

1. 386

/organism="Rattus norvegicus"

/mol\_type="mRNA"

/db\_xref="taxon:10116"

/clone="srp92-00054-g9"

/tissue\_type="pituitary gland brain"

/clone\_lib="srp92 (10238)"

/note="Vector: pSP0rt1; Site\_1: SalI; Site\_2: NotI; peneal  
gland brain region"

BASE COUNT 57 a 149 c 119 g 60 t 1 others

ORIGIN

Query Match 42.1%; Score 27.8; DB 14; Length 386;  
Best Local Similarity 64.1%; Pred. No. 48;  
Matches 41; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

QY 2 TGAAGACAGCTATCGGATTCAGTGGCACTGGCTGGTTTCGCTACCGTGCCCGGCGG 61

Db 85 TGAAGACAGCTATCGGATTCAGTGGCACTGGCTGGTTTCGCTACCGTGCCCGGCGG 26

QY 62 CGGC 65

Db 25 CGGC 22

RESULT 3

AI635506/c

LOCUS

DEFINITION

ts95a03.x1 NCI\_CGAP\_GC6 Homo sapiens cDNA clone IMAGE:2238988 3',  
mRNA sequence.

ACCESSION AI635506

VERSION AI635506.1

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov

Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael

R. Emmert-Buck, M.D., Ph.D.

cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima

Bonaldo, Ph.D.

cDNA Library Arrayed by: Greg Lennon, Ph.D.

DNA Sequencing by: Washington University Genome Sequencing Center

Clone distribution: NCI-CGAP clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

www-bio.llnl.gov/bbrp/image/image.html

Insert length: 383 Std Error: 0.00

Seq primer: -40UP from Gibco.

Location/Qualifiers

1. 310

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clones="IMAGE:2238988"

/tissue\_type="pooled germ cell tumors"

/lab\_host="DH10B"

/clone\_lib="NCI CGAP GC6"

/notes="Vector: p7713D-Pac (Pharmacia) with a modified

polylinker; Site\_1: Not I; Site\_2: Eco RI; Plasmid DNA

from the normalized library NCI\_CGAP GC4 was prepared, and

ss circles were made in vitro. Following HAP purification,

this DNA was used as tracer in a subtractive hybridization

reaction. The driver was PCR-amplified cDNAs from a pool

of 5,000 clones made from the same library (cloneIDs

1257096-1258631, 1469084-1470983, and 1475592-1476743).

Subtraction by Bento Soares and M. Fatima Bonaldo.

BASE COUNT 71 a 72 c 64 g 103 t

ORIGIN

Query Match 40.6%; Score 26.8; DB 9; Length 310;

Best Local Similarity 64.5%; Pred. No. 97;

Matches 40; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

QY 5 AAAAGACAGCTATCGGATTCAGTGGCACTGGCTGGTTTCGCTACCGTGCCCGGCGG 64

Db 268 AAAAGACAGCTATCGGATTCAGTGGCACTGGCTGGTTTCGCTACCGTGCCCGGCGG 209

QY 65 CC 66

Db 208 CC 207

RESULT 4

AW004057/c

LOCUS

DEFINITION

wg85e09.x1 NCI\_CGAP\_GC6 Homo sapiens cDNA clone IMAGE:2478856 3',  
mRNA sequence.

ACCESSION AW004057

VERSION AW004057.1

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM

Homo sapiens



## FEATURES

source

## Location/Qualifiers

```

1. .325
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="GDB:5923670"
/db_xref="taxon:9606"
/clone="IMAGE:726760"
/sex="male"
/lab host="DH10B"
/clone_lib="Soares testis_NHT"
/note="Vector: pVT3D-Pac (Pharmacia) with a modified
polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
was prepared from mRNA obtained from Clontech Laboratories
, Inc., and primed with a Not I - oligo(dT) primer [5',
TGTTACCAATCTGAAGTGGAGCGCGCCCAATTTTTTTTTTTT 3'].
Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pVT73 vector. Library
went through one round of normalization to Cot5, and was
constructed by Bento Soares and M. Fatima Bonaldo. "
72 a 72 c 64 g 117 t
BASE COUNT
ORIGIN

Query Match 40.6%; Score 26.8; DB 9; Length 325;
Best Local Similarity 64.5%; Pred. No. 98;
Matches 40; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

Qy 5 AAAAGACAGCTATCGGATTCAGTGGCACTGGCTGCTTGTGCTACCGTGGCCCGGCGG 64
Db 282 AAAAAGACGATCATCCAGGCGCATTCACGCTTGTGCTTTGTTAGCAGAGCCCGGCGAG 223

Qy 65 CC 66
Db 222 CC 221

RESULT 7
AA401699
LOCUS
DEFINITION zt60e09.r1 Soares_testis_NHT Homo sapiens cDNA clone IMAGE:726760
5', mRNA sequence.
ACCESSION AA401699
VERSION AA401699.1 GI:2057290
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 340)
Hillier, L., Allen, M., Bowles, L., Dubouque, T., Geisler, G., Jost, S.,
Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin, J., Moore, B.,
Schellenberg, K., Steptoe, M., Tan, F., Theising, B., White, Y., Wylie,
T., Waterston, R., and Wilson, R.
WASHU-Merck EST Project 1997
Unpublished
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 392 Sd Error: 0.00
Seq primer: -28ml3 rev2 Et from Amersham.
Location/Qualifiers
1. .340
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="GDB:5923670"
/db_xref="taxon:9606"
/clone="IMAGE:726760"
/sex="male"

```

## FEATURES

source

## Location/Qualifiers

```

/clone_lib="Soares testis_NHT"
/note="Vector: pVT3D-Pac (Pharmacia) with a modified
polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
was prepared from mRNA obtained from Clontech Laboratories
, Inc., and primed with a Not I - oligo(dT) primer [5',
TGTTACCAATCTGAAGTGGAGCGCGCCCAATTTTTTTTTTTT 3'].
Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pVT73 vector. Library
went through one round of normalization to Cot5, and was
constructed by Bento Soares and M. Fatima Bonaldo. "
106 a 75 c 82 g 77 t
BASE COUNT
ORIGIN

Query Match 40.6%; Score 26.8; DB 9; Length 340;
Best Local Similarity 64.5%; Pred. No. 1e+02;
Matches 40; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

Qy 5 AAAAGACAGCTATCGGATTCAGTGGCACTGGCTGCTTGTGCTACCGTGGCCCGGCGG 64
Db 76 AAAAAGACGATCATCCAGGCGCATTCACGCTTGTGCTTTGTTAGCAGAGCCCGGCGAG 135

Qy 65 CC 66
Db 136 CC 137

RESULT 8
BY641012/c
LOCUS
DEFINITION BY641012 RIKEN full-length enriched, visual cortex Mus musculus
cDNA clone K430341N20 3', mRNA sequence.
ACCESSION BY641012
VERSION BY641012.1 GI:26976194
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 409)
Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S.,
Nikaido, I., Osato, N., Saito, R., Suzuki, H., Yamanaka, I., Kiyosawa, H.,
Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Schonbach, C.,
Gojobori, T., Baldarelli, R., Hill, D. P., Bult, C., Hume, D. A.,
Quackenbush, J., Schriml, L. M., Kanapin, A., Mateuda, H., Batalov, S.,
Beisel, K. W., Blake, J. A., Bradt, D., Brusic, V., Chothia, C., Corbani,
L. E., Cousins, S., Dalla, E., Dragani, T. A., Fletcher, C. F., Forrest,
A., Frazer, K. S., Gaasterland, T., Gariboldi, M., Gissi, C., Godzik, A.,
Gough, J., Grimmond, S., Gustincich, S., Hirokawa, N., Jackson, I. J.,
Jarvis, E. D., Kanai, A., Kawaji, H., Kawasawa, Y., Kedzierski, R. M.,
King, B. L., Kongaya, A., Kurochkin, I. V., Lee, Y., Lenhard, B., Lyons,
P. A., Maglott, D. R., Maltais, L., Marchionni, L., McKenzie, L., Mikki,
H., Nagashima, T., Numata, K., Okido, T., Pavan, W. J., Pertea, G.,
Pesole, G., Petrovsky, N., Pillai, R., Pontius, J. U., Qi, D.,
Ramachandran, S., Ravasi, T., Reed, J. C., Reed, D. J., Reid, J., Ring,
B. Z., Ringwald, M., Sandelin, A., Schneider, C., Semple, C. A., Setou,
M., Shimada, K., Sultana, R., Takenaka, Y., Taylor, M. S., Teasdale,
R. D., Tomita, M., Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y.,
Watanabe, Y., Wells, C., Wilming, L. G., Wynshaw-Boris, A., Yanagisawa,
M., Yang, I., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A.,
Carninci, P., Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura,
M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K.,
Arakawa, T., Fukuda, S., Hara, A., Hashizume, M., Imotani, K., Ishii,
Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata,
K., Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander,
E. S., Rogers, J., Birney, E. and Hayashizaki, Y.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
JOURNAL
MEDLINE
PUBMED
12466851
Contact: Yoshihide Hayashizaki

```

Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center(GSC), Yokohama Institute  
The Institute of Physical and Chemical Research (RIKEN)  
1-7-22 Sushiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
Tel: 81-45-503-9222  
Fax: 81-45-503-9216  
Email: genome-res@gsc.riken.go.jp,  
URL: http://genome.gsc.riken.go.jp/  
T., Imotani, K., Ishii, Y., Itoh, M., Kawai, J., Konno, H., Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Numazaki, R., Ohno, M., Sakai, K., Sakazume, N., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M., Waki, K., Watahiki, A., Muramatsu, M. and Hayashizaki, Y. Direct Submission

Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)  
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)  
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)  
Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)  
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.  
Tissues were provided by Michela Fagiolini and Takao K. Hensch (Laboratory for Neuronal Circuit Development Brain Science Institute RIKEN 2-1 Hiroosawa,Wako-shi,Saitama 351-0198 Japan ) whose assistance we gratefully acknowledge.  
Please visit our web site (http://genome.gsc.riken.go.jp) for further details.

#### FEATURES

source  
Location/Qualifiers  
1. 409  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="K430341N20"  
/tissue\_type="visual cortex"  
/clone\_lib="RIKEN full-length enriched, visual cortex"  
89 a 96 c 87 g 135 t 2 others

Query Match 40.6%; Score 26.8; DB 13; Length 409;  
Best Local Similarity 64.5%; Pred. No. 1.1e+02;  
Matches 40; Conservative 0; Mismatches 22; Indels 0; Gaps 0;  
Qy 1 ATGAAGACAGCTATCGCATTCGACGTGGCACTGGCTGCTTTCGCTACCGTGGCCGAG 60  
|||||  
Db 102 AGGAAAGAGGCAAGCAGGCTTAGAGAGGTGGCTGGCTTCTCAAGCTTCCACAG 43  
|||||  
Qy 61 GC 62  
||  
Db 42 GC 41

RESULT 9  
BU783334  
LOCUS  
DEFINITION in02905.y1 Human insulinoma Homo sapiens cDNA clone IMAGE:6123441  
5', mRNA sequence.  
ACCESSION BU783334  
VERSION BU783334.1 GI:23827413  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

#### REFERENCE

AUTHORS  
1 (bases 1 to 569)  
Melton, D., Brown, J., Kenty, G., Permutt, A., Lee, C., Kaestner, K., Lemishka, I., Searce, M., Brestelli, J., Gradwohl, G., Clifton, S., Hillier, L., Marra, M., Pape, D., Wylie, T., Martin, J., Blistain, A., Schmitt, A., Theising, B., Ritter, E., Ronko, I., Bennett, J., Cardenas, M., Gibbons, M., McCann, R., Cole, R., Tsagarishvili, R., Williams, T., Jackson, Y. and Bowers, Y.  
Endocrine Pancreas Consortium  
Unpublished  
Other ESTs: in02905.x1  
Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue  
Endocrine Pancreas Consortium  
Harvard University, Howard Hughes Medical Institute  
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge, MA 02138  
Tel: 617-495-1812  
Fax: 617-495-8557  
Email: dmelton@bioh.harvard.edu  
Library was constructed by Dr. J. Ferrer In vivo mass-excised to pBluescript SK- by Dr. H. Inoue DNA sequencing by: Washington University Genome Sequencing Center For information on obtaining a clone please contact: Dr. Hiroshi Inoue (hinoue@im.wustl.edu)  
Seq primer: -40RP from Gibco  
High quality sequence stop: 429.

#### TITLE

JOURNAL  
COMMENT

#### FEATURES

source  
Location/Qualifiers  
1. 569  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:6123441"  
/tissue\_type="insulinoma"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="Human insulinoma"  
/notes="Organ: pancreas; Vector: pBluescript SK-; Site: 1: XhoI; Site 2: EcoRI; Constructed with lambda ZapII system (Stratagene) by Dr. J. Ferrer, in vivo mass-excised to pBluescript SK- by Dr. H. Inoue following the Washington University protocol (http://genome.wustl.edu/est/lambda\_protocol.shtml). Please contact Hiroshi Inoue, MD/PhD for further information on this library (Metabolism Division, Permutt Laboratory, Washington University School of Medicine, Box 8127, 660 S Euclid Ave, St. Louis, MO 63110). Note: this is a Washington University Pancreas EST project library."  
118 a 160 c 144 g 147 t

Query Match 40.6%; Score 26.8; DB 13; Length 569;  
Best Local Similarity 64.5%; Pred. No. 1.2e+02;  
Matches 40; Conservative 0; Mismatches 22; Indels 0; Gaps 0;  
Qy 5 AAAAGACAGCTATCGCATTCGACGTGGCACTGGCTGCTTTCGCTACCGTGGCCGCGG 64  
|||||  
Db 253 AAAAAGCAGCATCACCAGGCCATTCACGCTTTGTGCTTTGTAGCAGGCCGCGAG 312  
|||||  
Qy 65 CC 66  
||  
Db 313 CC 314

RESULT 10  
BX099539  
LOCUS  
DEFINITION BX099539 Soares testis NHT Homo sapiens cDNA clone IMAGE:998231928  
; IMAGE:782638, mRNA sequence.  
ACCESSION BX099539  
VERSION BX099539.1 GI:27844114  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 667)





```
/clone lib="NIH MGC 82"
```

was gel purified and cloned into the dephospho

```

BASE COUNT      60 a      63 c      84 g      61 t
ORIGIN

Query Match      40.3%; Score 26.6; DB 28; Length 268;
Best Local Similarity 71.4%; Pred. No. 1.1e+02;
Matches 35; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 16 ATCGCGATTGCAGTGCCTGCTGTTTCGCTACCGTGGCCGACGCGG 64
    |||||
Db 189 ATCGTGTGCTGTGGCGGTCGCTTGGCTCCCGTCGCTCGCGG 141

RESULT 15
AU244728
LOCUS AU244728 Shibata Xenopus AEM lambda-ZAP II cDNA library Xenopus
DEFINITION laevis cDNA clone p5el2 5', mRNA sequence.
ACCESSION AU244728
VERSION AU244728.1 GI:18850654
KEYWORDS EST.
SOURCE Xenopus laevis (African clawed frog)
ORGANISM Xenopus laevis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
Xenopodinae; Xenopus.
REFERENCE 1 (bases 1 to 359)
AUTHORS Shibata,M., Itoh,M., Ohmori,S., Shinga,J. and Taira,M.
TITLE Systematic screening and expression analysis of the head organizer
JOURNAL genes in Xenopus embryos
MEDLINE Dev. Biol. 239 (2), 241-256 (2001)
PUBMED 21643879
COMMENT 11784032
Contact: Masanori Taira
Department of Biological Sciences
Graduate School of Science, University of Tokyo; CREST, Japan
Science and Technology Corporation, Japan
7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan
Tel: 81-03-5841-4434
Fax: 81-03-5841-4434
Email: m_taira@biol.s.u-tokyo.ac.jp,
URL:http://www.biol.s.u-tokyo.ac.jp/users/lmb/lmb-hp.html.
Location/Qualifiers
1..359
/organism="Xenopus laevis"
/mol_type="mRNA"
/db_xref="taxon:8355"
/clone="p5el2"
/tissue_type="anterior endomesoderm"
/dev_stage="late gastrula to early neurula stages 12.5-13"
/clone_lib="Shibata Xenopus AEM lambda-ZAP II cDNA
library"

BASE COUNT      90 a      92 c      82 g      93 t      2 others
ORIGIN

Query Match      40.3%; Score 26.6; DB 9; Length 359;
Best Local Similarity 63.1%; Pred. No. 1.2e+02;
Matches 41; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY 2 TGAAGAGACAGCTATCGGATTCGAGTGGCAGTGGTTCGCTACCGTGGCCGAGG 61
    |||||
Db 129 TGAAGAGGAGATATCGTATCGTATCGGTCATCGCTGGTTCGTCTCCAGCAGG 188

QY 62 CGGCC 66
    |||||
Db 189 CAGCC 193

```

Search completed: August 19, 2003, 08:29:16  
 Job time : 201.082 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 18, 2003, 19:53:48 ; Search time 4736.31 Seconds  
(without alignments)  
9743.047 Million cell updates/sec

Title: US-09-987-455-5  
Perfect score: 1128  
Sequence: 1 atgaaaagacagtcgtgc.....ttctgacaaatcgacg 1128

Scoring table: IDENTITY NUC  
Gapop 10.0, Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : GenEmbl.\*

- 1: gb.ba.\*
- 2: gb.htg.\*
- 3: gb.in.\*
- 4: gb.in.\*
- 5: gb.ov.\*
- 6: gb.pat.\*
- 7: gb.ph.\*
- 8: gb.pl.\*
- 9: gb.pr.\*
- 10: gb.ro.\*
- 11: gb.ats.\*
- 12: gb.sy.\*
- 13: gb.un.\*
- 14: gb.vi.\*
- 15: em.ba.\*
- 16: em.fun.\*
- 17: em.hum.\*
- 18: em.in.\*
- 19: em.mu.\*
- 20: em.om.\*
- 21: em.or.\*
- 22: em.ov.\*
- 23: em.pat.\*
- 24: em.ph.\*
- 25: em.pl.\*
- 26: em.ro.\*
- 27: em.sts.\*
- 28: em.un.\*
- 29: em.vi.\*
- 30: em.htg.hum.\*
- 31: em.htg.inv.\*
- 32: em.htg.other.\*
- 33: em.htg.mus.\*
- 34: em.htg.pln.\*
- 35: em.htg.rod.\*
- 36: em.htg.mam.\*
- 37: em.htg.vrt.\*
- 38: em.sy.\*
- 39: em.htgo.hum.\*
- 40: em.htgo.mus.\*
- 41: em.htgo.other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1128	100.0	1128	6	AX431268 Sequence
2	1128	100.0	1128	6	AX431271 Sequence
3	1065	94.4	1314	6	A27435 DNA sequenc
4	1065	94.4	1314	6	E01937 Synthetic D
5	1065	94.4	1341	6	A30593 DNA for tis
6	1065	94.4	1419	6	A27725 DNA sequenc
7	1065	94.4	1419	6	AR059996 Sequence
8	1065	94.4	1419	6	E01944 Synthetic D
9	1065	94.4	1689	6	I06609 Sequence 4
10	1065	94.4	1689	6	I06614 Sequence 46
11	1065	94.4	1689	6	I08789 Sequence 3
12	1065	94.4	1780	6	A06611 Synthetic n
13	1065	94.4	1780	6	A10226 Synthetic D
14	1065	94.4	1836	6	E01176 DNA encodin
15	1065	94.4	1836	6	I01581 Sequence 1
16	1065	94.4	1981	6	I01583 Sequence 3
17	1065	94.4	1989	6	I01256 Sequence 2
18	1065	94.4	2099	6	A07207 Artificial
19	1065	94.4	2100	6	A07197 H.sapiens m
20	1065	94.4	2100	6	A27429 cDNA sequen
21	1065	94.4	2100	6	E01934 DNA encodin
22	1065	94.4	2101	6	AR059986 Sequence
23	1065	94.4	2151	6	E00654 cDNA encodi
24	1065	94.4	2162	6	AR017909 Sequence
25	1065	94.4	2162	6	AR017931 Sequence
26	1065	94.4	2162	6	I18796 Sequence 3
27	1065	94.4	2249	6	A07209 Artificial
28	1065	94.4	2544	6	I08530 Sequence 1
29	1065	94.4	2547	6	E01163 cDNA encodi
30	1065	94.4	2547	6	I07841 Sequence 1
31	1065	94.4	2547	6	I07989 Sequence 1
32	1065	94.4	2547	6	I09622 Sequence 1
33	1065	94.4	2549	6	E01055 DNA encodin
34	1065	94.4	2560	6	A01465 H.sapiens t
35	1065	94.4	7533	6	I08638 Sequence 2
36	1064	94.3	1689	6	I08637 Sequence 1
37	1063.4	94.3	1314	6	AR059989 Sequence
38	1063.4	94.3	1506	6	E02247 cDNA sequen
39	1063.4	94.3	1689	6	E02247 tPA-DNA fro
40	1063.4	94.3	1689	6	A14927 tPA-DNA fro
41	1063.4	94.3	1689	6	A14930 tPA-DNA fro
42	1063.4	94.3	1689	6	A34204 Human t-PA
43	1063.4	94.3	1689	6	A34205 Human t-PA
44	1063.4	94.3	1689	6	AX427761 Sequence
45	1063.4	94.3	1689	6	E02448 DNA encodin
					E02663 DNA encodin

ALIGNMENTS

RESULT 1  
AX431268 LOCUS  
DEFINITION Sequence 2 from Patent WO0240650.  
ACCESSION AX431268  
VERSION AX431268.1 GI:21656150  
KEYWORDS synthetic construct  
SOURCE synthetic construct  
ORGANISM artificial sequences.  
REFERENCE 1  
AUTHORS Goetz,F., Werner,R.G., Tayapiwatana,C., Manosroi,A. and Manosroi,J.  
TITLE Methods for large scale production of recombinant dna-derived tpa  
or k2s molecules  
JOURNAL Patent: WO 0240650-A 2 23-MAY-2002;

```
BOEHRINGER INGELHEIM INT (DE)
FEATURES
  Location/Qualifiers
    1..1128
    /organism="synthetic construct"
    /mol_type="genomic DNA"
    /db_xref="taxon:32630"
    /note="coding sequence for OmpA-K2S fusion protein"
BASE COUNT      238 a 332 c 332 g 226 t
ORIGIN
Query Match      100.0%; Score 1128; DB 6; Length 1128;
Best Local Similarity 100.0%; Pred. No. 5.5e-259;
Matches 1128; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGAAAAGACAGCTATCGGATTCAGTGGCAGCTGGCTGGTTCGCTACCGTGGCCAG 60
DB 1 ATGAAAAGACAGCTATCGGATTCAGTGGCAGCTGGCTGGTTCGCTACCGTGGCCAG 60
QY 61 GCGGCTCTGAGGGAACAGTGAATCTTTGGGAATGGGTGAGCTACCGTGGCCAG 120
DB 61 GCGGCTCTGAGGGAACAGTGAATCTTTGGGAATGGGTGAGCTACCGTGGCCAG 120
QY 121 CACAGCTCAGCAGTGGGTGCTCTGCTCCCGTGGGAATTCATGATCCTGATAGGC 180
DB 121 CACAGCTCAGCAGTGGGTGCTCTGCTCCCGTGGGAATTCATGATCCTGATAGGC 180
QY 181 AAGGTTTACAGACAGAACCCAGTCCAGGCACTGGGCTGGGCAACATAATTAC 240
DB 181 AAGGTTTACAGACAGAACCCAGTCCAGGCACTGGGCTGGGCAACATAATTAC 240
QY 241 TGCCGGAATCCTGATGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCAGGCTG 300
DB 241 TGCCGGAATCCTGATGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCAGGCTG 300
QY 301 ACGTGGAGTACTGTGATGTGCTCTGCTCCAGCTCGGCTGAGACAGTACAGCCAG 360
DB 301 ACGTGGAGTACTGTGATGTGCTCTGCTCCAGCTCGGCTGAGACAGTACAGCCAG 360
QY 361 CCTCAGTTTCCATCAAGAGGAGGCTCTTCGCCGACATCGCTCCACCCCTGGCAGGCT 420
DB 361 CCTCAGTTTCCATCAAGAGGAGGCTCTTCGCCGACATCGCTCCACCCCTGGCAGGCT 420
QY 421 GCCATCTTTGCCAAGCAGGAGGTGCGCCCGAGAGCGGTTCTGTGCGGGGGCATATCTC 480
DB 421 GCCATCTTTGCCAAGCAGGAGGTGCGCCCGAGAGCGGTTCTGTGCGGGGGCATATCTC 480
QY 481 ATCAGCTCCTGTGATGTTCTCTGCGGCCCACTGCTTCCAGAGAGGTTTCCGCCCCAC 540
DB 481 ATCAGCTCCTGTGATGTTCTCTGCGGCCCACTGCTTCCAGAGAGGTTTCCGCCCCAC 540
QY 541 CACCTGACGGTGTATCTTGGGAGAACATACCGGGTGGTCCCTGGCGAGGAGGAGCAGAAA 600
DB 541 CACCTGACGGTGTATCTTGGGAGAACATACCGGGTGGTCCCTGGCGAGGAGGAGCAGAAA 600
QY 601 TTTGAAGTCGAAAATACATTTGCCATAAGGAATTCGATGATGACACTTACGACAAATGAC 660
DB 601 TTTGAAGTCGAAAATACATTTGCCATAAGGAATTCGATGATGACACTTACGACAAATGAC 660
QY 661 ATTGGGCTGTGACGTGAATTCGATGTTGCTCCCGCTGTGCCAGGAGAGAGCGTGGTC 720
DB 661 ATTGGGCTGTGACGTGAATTCGATGTTGCTCCCGCTGTGCCAGGAGAGAGCGTGGTC 720
QY 721 CGCAGTGTGCTTTCCCGCGGAGCTGCGAGCTGCGGAGCTGGAGCGAGTGTGAGCTC 780
DB 721 CGCAGTGTGCTTTCCCGCGGAGCTGCGAGCTGCGGAGCTGGAGCGAGTGTGAGCTC 780
QY 781 TCCGCTACGGCAAGCATGAGGCTTGTCTCTTTCTATTTCGAGCGGCTGAAGAGGCT 840
DB 781 TCCGCTACGGCAAGCATGAGGCTTGTCTCTTTCTATTTCGAGCGGCTGAAGAGGCT 840
QY 841 CATGTACAGTGTACCCATCCAGCGGCTGACATCAGCAATTTACTTAAACAGACAGTC 900
DB 841 CATGTACAGTGTACCCATCCAGCGGCTGACATCAGCAATTTACTTAAACAGACAGTC 900

RESULT 2
AX431271
LOCUS      AX431271      1128 bp      DNA      linear      PAT 28-JUN-2002
DEFINITION Sequence 5 from Patent WO0240650.
ACCESSION  AX431271
VERSION     AX431271.1  GI:21656153
KEYWORDS
SOURCE      synthetic construct
            synthetic construct
            artificial sequences.
REFERENCE   1
AUTHORS     Getz, F., Werner, R.G., Tayapiwatana, C., Manosroi, A. and Manosroi, J.
TITLE       Methods for large scale production of recombinant dna-derived tpa
            or k2s molecules
            Patent: WO 0240650-A 5 23-MAY-2002;
            BOEHRINGER INGELHEIM INT. (DE)
JOURNAL
FEATURES
  source
    1..1128
    /organism="synthetic construct"
    /mol_type="genomic DNA"
    /db_xref="taxon:32630"
    /note="coding sequence for OmpA-K2S fusion protein"
BASE COUNT      238 a 332 c 332 g 226 t
ORIGIN
Query Match      100.0%; Score 1128; DB 6; Length 1128;
Best Local Similarity 100.0%; Pred. No. 5.5e-259;
Matches 1128; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGAAAAGACAGCTATCGGATTCAGTGGCAGCTGGCTGGTTCGCTACCGTGGCCAG 60
DB 1 ATGAAAAGACAGCTATCGGATTCAGTGGCAGCTGGCTGGTTCGCTACCGTGGCCAG 60
QY 61 GCGGCTCTGAGGGAACAGTGAATCTTTGGGAATGGGTGAGCTACCGTGGCCAG 120
DB 61 GCGGCTCTGAGGGAACAGTGAATCTTTGGGAATGGGTGAGCTACCGTGGCCAG 120
QY 121 CACAGCTCAGCAGTGGGTGCTCTGCTCCCGTGGGAATTCATGATCCTGATAGGC 180
DB 121 CACAGCTCAGCAGTGGGTGCTCTGCTCCCGTGGGAATTCATGATCCTGATAGGC 180
QY 181 AAGGTTTACAGACAGAACCCAGTCCAGGCACTGGGCTGGGCAACATAATTAC 240
DB 181 AAGGTTTACAGACAGAACCCAGTCCAGGCACTGGGCTGGGCAACATAATTAC 240
QY 241 TGCCGGAATCCTGATGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCAGGCTG 300
DB 241 TGCCGGAATCCTGATGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCAGGCTG 300
QY 301 ACGTGGAGTACTGTGATGTGCTCTGCTCCAGCTCGGCTGAGACAGTACAGCCAG 360
DB 301 ACGTGGAGTACTGTGATGTGCTCTGCTCCAGCTCGGCTGAGACAGTACAGCCAG 360
QY 361 CCTCAGTTTCCATCAAGAGGAGGCTCTTCGCCGACATCGCTCCACCCCTGGCAGGCT 420
DB 361 CCTCAGTTTCCATCAAGAGGAGGCTCTTCGCCGACATCGCTCCACCCCTGGCAGGCT 420
QY 421 GCCATCTTTGCCAAGCAGGAGGTGCGCCCGAGAGCGGTTCTGTGCGGGGGCATATCTC 480
DB 421 GCCATCTTTGCCAAGCAGGAGGTGCGCCCGAGAGCGGTTCTGTGCGGGGGCATATCTC 480
QY 481 ATCAGCTCCTGTGATGTTCTCTGCGGCCCACTGCTTCCAGAGAGGTTTCCGCCCCAC 540
DB 481 ATCAGCTCCTGTGATGTTCTCTGCGGCCCACTGCTTCCAGAGAGGTTTCCGCCCCAC 540
QY 541 CACCTGACGGTGTATCTTGGGAGAACATACCGGGTGGTCCCTGGCGAGGAGGAGCAGAAA 600
DB 541 CACCTGACGGTGTATCTTGGGAGAACATACCGGGTGGTCCCTGGCGAGGAGGAGCAGAAA 600
QY 601 TTTGAAGTCGAAAATACATTTGCCATAAGGAATTCGATGATGACACTTACGACAAATGAC 660
DB 601 TTTGAAGTCGAAAATACATTTGCCATAAGGAATTCGATGATGACACTTACGACAAATGAC 660
QY 661 ATTGGGCTGTGACGTGAATTCGATGTTGCTCCCGCTGTGCCAGGAGAGAGCGTGGTC 720
DB 661 ATTGGGCTGTGACGTGAATTCGATGTTGCTCCCGCTGTGCCAGGAGAGAGCGTGGTC 720
QY 721 CGCAGTGTGCTTTCCCGCGGAGCTGCGAGCTGCGGAGCTGGAGCGAGTGTGAGCTC 780
DB 721 CGCAGTGTGCTTTCCCGCGGAGCTGCGAGCTGCGGAGCTGGAGCGAGTGTGAGCTC 780
QY 781 TCCGCTACGGCAAGCATGAGGCTTGTCTCTTTCTATTTCGAGCGGCTGAAGAGGCT 840
DB 781 TCCGCTACGGCAAGCATGAGGCTTGTCTCTTTCTATTTCGAGCGGCTGAAGAGGCT 840
QY 841 CATGTACAGTGTACCCATCCAGCGGCTGACATCAGCAATTTACTTAAACAGACAGTC 900
DB 841 CATGTACAGTGTACCCATCCAGCGGCTGACATCAGCAATTTACTTAAACAGACAGTC 900
```

Db 361 CCTCAGTTTCGCATCAAGAGGGGCTCTTCGCGGACATCGCTCCACCCCTGGGAGGCT 420  
Qy 421 GCCATCTTTGCCAAGACAGGAGGTCGCCCGAGAGCGGTTCTCTGTGCGGGGCATATCTC 480  
Db 421 GCCATCTTTGCCAAGACAGGAGGTCGCCCGAGAGCGGTTCTCTGTGCGGGGCATATCTC 480  
Qy 481 ATCAGCTCTCTGATGATCTCTCTGCGCGCCACTCTCTTCAGGAGAGGTTTCCGCGCCAC 540  
Db 481 ATCAGCTCTCTGATGATCTCTCTGCGCGCCACTCTCTTCAGGAGAGGTTTCCGCGCCAC 540  
Qy 541 CACCTGACGGTGATCTTGGGAGAAACATACCGGGTGGTCCCTGGGAGAGGAGCAGAAA 600  
Db 541 CACCTGACGGTGATCTTGGGAGAAACATACCGGGTGGTCCCTGGGAGAGGAGCAGAAA 600  
Qy 601 TTTGAAGTCGAAAAATAATTGTCTCAATAGGAATTCGATGATGACATTCACGAAATGAC 660  
Db 601 TTTGAAGTCGAAAAATAATTGTCTCAATAGGAATTCGATGATGACATTCACGAAATGAC 660  
Qy 661 ATTGGCTGCTGACGCTGAATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTC 720  
Db 661 ATTGGCTGCTGACGCTGAATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTC 720  
Qy 721 CGCAGTGTGTCTTCCCGCGGACCTGACGCTGCGGACTGCGGAGTGTGAGCTC 780  
Db 721 CGCAGTGTGTCTTCCCGCGGACCTGACGCTGCGGAGTGTGAGCTC 780  
Qy 781 TCCGCTACGCAAGCATGAGGCTTGTCTCTTTCTATTTCGAGCGGCTGAAGAGGCT 840  
Db 781 TCCGCTACGCAAGCATGAGGCTTGTCTCTTTCTATTTCGAGCGGCTGAAGAGGCT 840  
Qy 841 CATGTACAGCTGTACCCATCCAGCGCTGACATCAACATTTACTTTAAACAGAACATGTC 900  
Db 841 CATGTACAGCTGTACCCATCCAGCGCTGACATCAACATTTACTTTAAACAGAACATGTC 900  
Qy 901 ACCGACAACTGTGTGTGGAGACACTCGGAGCGGCGGCCCGAGCAAACTTGCAC 960  
Db 901 ACCGACAACTGTGTGTGGAGACACTCGGAGCGGCGGCCCGAGCAAACTTGCAC 960  
Qy 961 GACGCTGCCAGGGCGATTTCGGAGGCGCCCTGTGTGTCTGAAAGATGCGCGCATGACT 1020  
Db 961 GACGCTGCCAGGGCGATTTCGGAGGCGCCCTGTGTGTCTGAAAGATGCGCGCATGACT 1020  
Qy 1021 TTGGTGGGCATCATCAGCTGGGGCTTGGGCTGTGGAGAGAGATGTCGGGGTGTGTAC 1080  
Db 1021 TTGGTGGGCATCATCAGCTGGGGCTTGGGCTGTGGAGAGAGATGTCGGGGTGTGTAC 1080  
Qy 1081 ACAAAGTTTACCACTACCTAGACTGGATTTCGTGACACATGCGGACCG 1128  
Db 1081 ACAAAGTTTACCACTACCTAGACTGGATTTCGTGACACATGCGGACCG 1128

RESULT 3  
A27435  
LOCUS A27435 1314 bp DNA linear PAT 09-JUL-2002  
DEFINITION DNA sequence of coding region in pTKPdelta1atrp.  
ACCESSION A27435  
VERSION A27435.1 GI:21727230  
KEYWORDS  
SOURCE  
ORGANISM  
synthetic construct  
synthetic construct  
artificial sequences  
1. (bases 1 to 1314)  
Niwa,M., Saito,Y., Sasaki,H., Hayashi,M., Notani,J. and  
Kobayashi,M.  
TITLE New tissue plasminogen activator  
JOURNAL Patent: EP 0302456-A 35 08-FEB-1989;  
FUJISAWA PHARMACEUTICAL CO., LTD  
FEATURES  
Location/Qualifiers  
1. .1314  
/organism="synthetic construct"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:32630"  
BASE COUNT 286 a 386 c 393 g 249 t

ORIGIN  
Query Match 94.4%; Score 1065; DB 6; Length 1314;  
Best Local Similarity 99.5%; Pred. No. 6.5e-244;  
Matches 1068; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
Qy 56 CCCAGGCGGCTCTGAGGGAAACAGTACTGCTACTTTTGGGAATGGGTACGCTTACCCTG 115  
Db 239 CCCCTGCTGCTCTGAGGGAAACAGTACTGCTACTTTTGGGAATGGGTACGCTTACCCTG 298  
Qy 116 GCAGCGACAGCTCACCGAGTCCGGTGCCTCTCCCTCCGCTCCGCTGGAATTCATGATCTGA 175  
Db 299 GCAGCGACAGCTCACCGAGTCCGGTGCCTCTCCCTCCGCTCCGCTGGAATTCATGATCTGA 358  
Qy 176 TAGCAAGGTTTACACAGCACAGAAACCCAGTGCAGGCACTGGGCTGGGCAAAACATA 235  
Db 359 TAGCAAGGTTTACACAGCACAGAAACCCAGTGCAGGCACTGGGCTGGGCAAAACATA 418  
Qy 236 ATTAAGTCCGGAATCTGATGGGATGCAAGCCCTGGTGCACGCTGCTGAAGAACCGCA 295  
Db 419 ATTAAGTCCGGAATCTGATGGGATGCAAGCCCTGGTGCACGCTGCTGAAGAACCGCA 478  
Qy 296 GGCTGAGCTGGGAGTACTGTGATGTCCTCTCTCCCTCCGCTGAGACAGATACA 355  
Db 479 GGCTGAGCTGGGAGTACTGTGATGTCCTCTCTCCCTCCGCTGAGACAGATACA 538  
Qy 356 GCCAGCTCAGTTTCGCATCAAGAGAGGGCTCTTCGCGGACATCGCTCCACCCCTGGC 415  
Db 539 GCCAGCTCAGTTTCGCATCAAGAGAGGGCTCTTCGCGGACATCGCTCCACCCCTGGC 598  
Qy 416 AGGCTGCCATCTTTGCCAAGCACAGGAGGTGCGCCGAGAGCGGTTCTGTGCGGGGCA 475  
Db 599 AGGCTGCCATCTTTGCCAAGCACAGGAGGTGCGCCGAGAGCGGTTCTGTGCGGGGCA 658  
Qy 476 TACTCATCAGCTCTGTGATTTCTCTGCGCGGCACTGCTTCCAGAGAGGTTTCCGC 535  
Db 659 TACTCATCAGCTCTGTGATTTCTCTGCGCGGCACTGCTTCCAGAGAGGTTTCCGC 718  
Qy 536 CCACACACCTGACGGTGATCTTGGGAGAAACATACCGGGTGGTCCCTGGCGAGGAGGAGC 595  
Db 719 CCACACACCTGACGGTGATCTTGGGAGAAACATACCGGGTGGTCCCTGGCGAGGAGGAGC 778  
Qy 596 AGAAATTTGAAGTCGAAAAATACATTTCCATAAGGAATTCGATGATGACATTTACGACA 655  
Db 779 AGAAATTTGAAGTCGAAAAATACATTTCCATAAGGAATTCGATGATGACATTTACGACA 838  
Qy 656 ATGACATTTGCGCTGCTGACGCTGAAATCGAATTCGCTCCGCTGTCGCGAGAGAGAGCG 715  
Db 839 ATGACATTTGCGCTGCTGACGCTGAAATCGAATTCGCTCCGCTGTCGCGAGAGAGAGCG 898  
Qy 716 TGGTCCGCACTGTGTGCTTTCCCGCGGAGACCTGACAGCTGCGGAGCTGGAGCGAGTGTG 775  
Db 899 TGGTCCGCACTGTGTGCTTTCCCGCGGAGACCTGACAGCTGCGGAGCTGGAGCGAGTGTG 958  
Qy 776 AGCTCTCCGGCTACGGCAAGCATGAGGCTTGTCTCTCTTTCTATTTCGAGCGGCTGAAGG 835  
Db 959 AGCTCTCCGGCTACGGCAAGCATGAGGCTTGTCTCTCTTTCTATTTCGAGCGGCTGAAGG 1018  
Qy 836 AGGCTCATGTGACACTGTACCCATCCAGCGCTGCAATCACAAATTTACTTTAAACAGAA 895  
Db 1019 AGGCTCATGTGACACTGTACCCATCCAGCGCTGCAATCACAAATTTACTTTAAACAGAA 1078  
Qy 896 CAGTCACCGACACATGCTGTGTGCTGAGACACTCGGAGCGGCGGCGCCAGCAAACT 955  
Db 1079 CAGTCACCGACACATGCTGTGTGCTGAGACACTCGGAGCGGCGGCGCCAGCAAACT 1138  
Qy 956 TGACAGCAGCTCTGCGGAGGATTTCCGGAGGCGCCCTTGGTGTGTCTGAAAGATGGCCGCA 1015  
Db 1139 TGACAGCAGCTCTGCGGAGGCGATTCGGAGGCGCCCTTGGTGTGTCTGAAAGATGGCCGCA 1198  
Qy 1016 TGACTTTGGTGGGCATCATCAGCTGGGGCTTGGGCTGTGAGCAAGAGATGTCCCGGTG 1075  
Db 1199 TGACTTTGGTGGGCATCATCAGCTGGGGCTTGGGCTGTGAGCAAGAGATGTCCCGGTG 1258

QY 1076 TGTACAAAAGTTTACCACTACCTAGACTGGATTGCGTGAACAACATGCGACCG 1128  
 DB 1259 TGTACAAAAGTTTACCACTACCTAGACTGGATTGCGTGAACAACATGCGACCG 1311

RESULT 4  
 E01937  
 LOCUS 1314 bp DNA linear PAT 29-SEP-1997  
 DEFINITION Synthetic DNA encoding new t-PA tissue plasminogen activator.  
 ACCESSION E01937  
 VERSION E01937.1 GI:2170185  
 KEYWORDS JP 1989104167-A/4.  
 SOURCE synthetic construct  
 ORGANISM synthetic construct  
 artificial sequences.  
 REFERENCE 1 (bases 1 to 1314)  
 AUTHORS Niwa, M., Saito, Y., Sasaki, H., Hayashi, M., Notani, J. and Kobayashi, M.  
 TITLE NOVEL TISSUE PLASMINOGEN ACTIVATION FACTOR  
 JOURNAL Patent: JP 1989104167-A 4 21-APR-1989;  
 FUJISAWA PHARMACEUT CO LTD  
 COMMENT OS Artificial gene  
 OC Artificial sequence; Genes.  
 PD 21-APR-1989  
 PF 01-AUG-1988 JP 1988192320  
 PR 03-AUG-1987 GB 87 8718298, 26-OCT-1987 GB 87 8725052, PR 13-NOV-1987 GB 87 8726683  
 PI NIWA MINEO, SAITO YOSHIMASA, SASAKI HITOSHI, HAYASHI MASAKO, PI NOTANI JOJI,  
 PI KOBAYASHI MASAKAZU  
 PC C12N9/64, C12N1/20, C12N5/00, C12N15/00//A61K37/54, (C12N9/64, PC C12N1/19),  
 PC (C12N9/64, C12N1/91);  
 CC strandedness: Double;  
 CC topology: Linear;  
 CC hypothetical: No;  
 CC anti-sense: No;  
 CC \*source: clones-pTQKpMdelatrp;  
 FH Key Location/Qualifiers  
 FT CDS 1..1314  
 /product='t-PA'.  
 FEATURES source  
 1..1314  
 Location/Qualifiers  
 /organism="synthetic construct"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:32630"  
 BASE COUNT 286 a 386 c 393 g 249 t  
 ORIGIN

Query Match 94.4%; Score 1065; DB 6; Length 1314;  
 Best Local Similarity 99.5%; Pred. No. 6.5e-244;  
 Matches 1068; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 56 CCCAGCGGCTCTGAGGAAACAGTACTTGGGAATGGGTACCCCTACCGTG 115  
 DB 239 CCCCTGCTCTGAGGAAACAGTACTTGGGAATGGGTACCCCTACCGTG 298  
 QY 116 GCACGACAGCTCACCCAGTGGGTGCTCTGCTCCCGTGGAAATCCATGATCCTGA 175  
 DB 299 GCACGACAGCTCACCCAGTGGGTGCTCTGCTCCCGTGGAAATCCATGATCCTGA 358  
 QY 176 TAGGCAAGTTTACACGACAGAACCCAGTGGCCAGGCACTGGGCTGGGCAACATA 235  
 DB 359 TAGGCAAGTTTACACGACAGAACCCAGTGGCCAGGCACTGGGCTGGGCAACATA 418  
 QY 236 ATTACTCGGAATCTGATGGGATGCAAGCCCTGTGTGACATGCTGAAGACCGCA 295  
 DB 419 ATTACTCGGAATCTGATGGGATGCAAGCCCTGTGTGACATGCTGAAGACCGCA 478  
 QY 296 GGCTGACGTGGAGTACTGTGATGTGCTCTGCTCCACCTGGCGGCTGAGACAGTACA 355

DB 479 GGCTGACGTGGAGTACTGTGATGTGCCCTCTGCTCCACTGGCGCTGAGACAGTACA 538  
 QY 356 GCCAGCCTCAGTTTCGCATCAAGAGAGGCTCTTCGCCGACATGCCCTCCACCCCTGGC 415  
 DB 539 GCCAGCCTCAGTTTCGCATCAAGAGAGGCTCTTCGCCGACATGCCCTCCACCCCTGGC 598  
 QY 416 AGGCTGCCATCTTTGCCAAGCACAGGAGGTGCCCGGAGAGCGGTCTCTGCGGGGCA 475  
 DB 599 AGGCTGCCATCTTTGCCAAGCACAGGAGGTGCCCGGAGAGCGGTCTCTGCGGGGCA 658  
 QY 476 TACTCATCAGTCTCTGCTGGATTCTCTTCGCCGCCACTGCTTCAGGAGAGGTTTCCGC 535  
 DB 659 TACTCATCAGTCTCTGCTGGATTCTCTTCGCCGCCACTGCTTCAGGAGAGGTTTCCGC 718  
 QY 536 CCACACCACTGACGGTGATCTTGGGCAAGACATACCGGGTGGTCTTGGCGAGAGGAGC 595  
 DB 719 CCACACCACTGACGGTGATCTTGGGCAAGACATACCGGGTGGTCTTGGCGAGGAGGAGC 778  
 QY 596 AGAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA 655  
 DB 779 AGAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA 838  
 QY 656 ATGACATTGGCGTCTGCAGCTGAAATCGGATTTCGCTCCCGTGTGCCAGGAGAGCAGCG 715  
 DB 839 ATGACATTGGCGTCTGCAGCTGAAATCGGATTTCGCTCCCGTGTGCCAGGAGAGCAGCG 898  
 QY 716 TGGTCCGACATGTGTGCTTCCCGCGGACCTGACGTGCCGAGCTGAGCGAGTGTG 775  
 DB 899 TGGTCCGACATGTGTGCTTCCCGCGGACCTGACGTGCCGAGCTGGAACGAGTGTG 958  
 QY 776 AGCTCTCGGCTAGCGCAAGCAGTGGAGCTGTCTCTCTTCTATTTCGGAGCGGCTGAAGG 835  
 DB 959 AGCTCTCGGCTAGCGCAAGCAGTGGAGCTGTCTCTCTTCTATTTCGGAGCGGCTGAAGG 1018  
 QY 836 AGGCTCATGTGACACTGTACCCATCCAGCGCTGCACATCACAACTTACTTAACAGAA 895  
 DB 1019 AGGCTCATGTGACACTGTACCCATCCAGCGCTGCACATCACAACTTACTTAACAGAA 1078  
 QY 896 CAGTACACGACATCATCTGTGTGCTGAGACACTCGGAGCGGGGCGGCGGCAAACT 955  
 DB 1079 CAGTACACGACATCATCTGTGTGCTGAGACACTCGGAGCGGGGCGGCGGCAAACT 1138  
 QY 956 TGACAGACGCTGCGCAGGCGGATTCGGAGGCGCCCTGTGTGTCTGAACGATGCGCGCA 1015  
 DB 1139 TGACAGACGCTGCGCAGGCGGATTCGGAGGCGCCCTGTGTGTCTGAACGATGCGCGCA 1198  
 QY 1016 TGACTTTGTGGGCATCATCAGCTGGGGCCTTGGGCTGTGGACAGAAGGATGTCCCGGTG 1075  
 DB 1199 TGACTTTGTGGGCATCATCAGCTGGGGCCTTGGGCTGTGGACAGAAGGATGTCCCGGTG 1258  
 QY 1076 TGACACAAAGTTTACCACTACCTAGACTGGATTCTGTGACAACTGCGACCG 1128  
 DB 1259 TGTACACAAAGTTTACCACTACCTAGACTGGATTCTGTGACAACTGCGACCG 1311

RESULT 5  
 A30593  
 LOCUS 1341 bp DNA linear PAT 04-AUG-1995  
 DEFINITION DNA for tissue plasminogen activator (t-PA) from patent EP0400545.  
 ACCESSION A30593  
 VERSION A30593.1 GI:1249165  
 KEYWORDS synthetic construct  
 SOURCE synthetic construct  
 ORGANISM synthetic construct  
 artificial sequences.  
 REFERENCE 1 (bases 1 to 1341)  
 AUTHORS Stern, A., Kohnert, U., Rudolph, R., Fischer, S. and Martin, U.  
 TITLE Derivative of tissue plasminogen activator  
 JOURNAL Patent: EP 0400545-A 2 05-DEC-1990;  
 BOEHRINGER MANNHEIM GMBH  
 FEATURES Location/Qualifiers  
 1..1341

		/organism="synthetic construct"								
		/mol_type="genomic DNA"								
		/db_xref="taxon:32630"								
BASE COUNT	293 a	397 c	399 g	252 t						
ORIGIN										
Query Match										
Best Local Similarity 94.4%; Score 1065; DB 6; Length 1341;										
Matches 1068; Conservative 0; Mismatches 5; Indels 0; Gaps 0;										
Qy	56	CCCAGCGCGCTCTGAGGAAACAGTACTGCTACTTTGGGAATGGGTGAGCTACCGTG	115							
Db	269	CCCCGCTGCTCTGAGGAAACAGTACTGCTACTTTGGGAATGGGTGAGCTACCGTG	328							
Qy	116	GCACGACAGCTTACCCAGTGGGTGCTCTCTGCTCCCGTGGAAATCCATGATCTGA	175							
Db	329	GCACGACAGCTTACCCAGTGGGTGCTCTCTGCTCCCGTGGAAATCCATGATCTGA	388							
Qy	176	TAGGCAAGGTTTACACACACAGACCCAGTGGTCCAGGCACTGGGCTGGGCAACATA	235							
Db	389	TAGGCAAGGTTTACACACACAGACCCAGTGGTCCAGGCACTGGGCTGGGCAACATA	448							
Qy	236	ATTACTGCGGAATCCTGATGGGATGCCAAGCCCTGGTGCCACGTCTGAAGAACCGCA	295							
Db	449	ATTACTGCGGAATCCTGATGGGATGCCAAGCCCTGGTGCCACGTCTGAAGAACCGCA	508							
Qy	296	GGTGACCTGGGAGTACTGTGATGTCCTCTCTGCTCCACCTGGCGCTGAGACAGTACA	355							
Db	509	GGTGACCTGGGAGTACTGTGATGTCCTCTCTGCTCCACCTGGCGCTGAGACAGTACA	568							
Qy	356	GCCAGCCTCAGTTTGGATCAAGAGGAGCTCTTGGCGGACATGCGCTCCACCCCTGGC	415							
Db	569	GCCAGCCTCAGTTTGGATCAAGAGGAGGCTCTTGGCGGACATGCGCTCCACCCCTGGC	628							
Qy	416	AGGTGCGCATCTTCCCAAGCACAGAGGTGCGCCGGAGAGCGGTTCCTGTCGGGGCA	475							
Db	629	AGGTGCGCATCTTCCCAAGCACAGAGGTGCGCCGGAGAGCGGTTCCTGTCGGGGCA	688							
Qy	476	TACTCATCAGCTCTCTGCTGGATTTCTCTGCGCGCCACTGCTTCAGGAGAGGTTTCGC	535							
Db	689	TACTCATCAGCTCTCTGCTGGATTTCTCTGCGCGCCACTGCTTCAGGAGAGGTTTCGC	748							
Qy	536	CCCACCCTGACCGGTGATCTTGGGCAACATACCGGGTGGTTCCTGGCGAGAGGAGC	595							
Db	749	CCCACCCTGACCGGTGATCTTGGGCAACATACCGGGTGGTTCCTGGCGAGGAGGC	808							
Qy	596	AGAAATTTGAAGTCGAAATATACATTTGCCATATAGGAATTCGATGACACTTACGACA	655							
Db	809	AGAAATTTGAAGTCGAAATATACATTTGCCATATAGGAATTCGATGACACTTACGACA	868							
Qy	656	ATGACATTCGCTGTGTCAGCTGAAATCGGATTCGTCCCGCTGTGCCCGAGAGCAGCG	715							
Db	869	ATGACATTCGCTGTGTCAGCTGAAATCGGATTCGTCCCGCTGTGCCCGAGAGCAGCG	928							
Qy	716	TGGTCCGCACTGTGTGCTTCCCGCGGACCTGTCAGCTCCGGAATGAGACGAGTGTG	775							
Db	929	TGGTCCGCACTGTGTGCTTCCCGCGGACCTGTCAGCTCCGGAATGAGACGAGTGTG	988							
Qy	776	AGCTCTCCGCTACGGCAAGCATGAGGCTTGTCTCTCTTCTATTCGGAGCGGCTGAAG	835							
Db	989	AGCTCTCCGCTACGGCAAGCATGAGGCTTGTCTCTCTTCTATTCGGAGCGGCTGAAG	1048							
Qy	836	AGGCTCATGTGACAGTGTACCCATCCAGCCGTGCACATCAACATTTACTTAACAGAA	895							
Db	1049	AGGCTCATGTGACAGTGTACCCATCCAGCCGTGCACATCAACATTTACTTAACAGAA	1108							
Qy	896	CAGTACCCGACAAATGTGTGTGTGTGGAGACATCTCGAGCGCGGCGGCCAGGCAACT	955							
Db	1109	CAGTACCCGACAAATGTGTGTGTGTGGAGACATCTCGAGCGCGGCGGCCAGGCAACT	1168							
Qy	956	TGCACAGACCTGCCAGGCGGATTCGGAGAGGCCCTCTGTGTGTGTAACATGGCGCA	1015							
Db	1169	TGCACAGCCTGCCAGGCGGATTCGGAGAGGCCCTCTGTGTGTGTAACATGGCGCA	1228							
RESULT 6										
LOCUS A27725 1419 bp DNA linear PAT 09-JUL-2002										
DEFINITION DNA sequence of coding region in pmTQk112.										
ACCESSION A27725										
VERSION A27725.1 GI:21727236										
KEYWORDS synthetic construct										
SOURCE synthetic construct										
ORGANISM synthetic construct										
REFERENCE 1 (bases 1 to 1419)										
AUTHORS Niwa,M., Saito,Y., Sasaki,H., Hayashi,M., Notani,J. and Kobayashi,M.										
TITLE New tissue plasminogen activator										
JOURNAL Patent: EP 0302456-A 49 08-FEB-1989;										
FUJISAWA PHARMACEUTICAL CO., LTD										
FEATURES										
source										
1..1419										
/organism="synthetic construct"										
/mol_type="genomic DNA"										
/db_xref="taxon:32630"										
BASE COUNT	309 a	413 c	426 g	271 t						
ORIGIN										
Query Match										
Best Local Similarity 94.4%; Score 1065; DB 6; Length 1419;										
Matches 1068; Conservative 0; Mismatches 5; Indels 0; Gaps 0;										
Qy	56	CCCAGCGCGCTCTCAGGAAACAGTACTGCTACTTTGGGAATGGGTGAGCTACCGTG	115							
Db	1229	TGACITTTGGTGGGCATCATCAGCTGGGCGCTGGGCTGTGGACAGAAGATGTCCCGGTTG	1288							
Qy	1076	TGTACACAAAGGTTACCAACTACCTAGACTGGATTCGTGACCAACATCGGACCG	1128							
Db	1289	TGTACACAAAGGTTACCAACTACCTAGACTGGATTCGTGACCAACATCGGACCG	1341							

QY 596 AGAAATTTGAAGTCCGAAATATCATTTGCTCCATAAGGAATTCGATGATGACACTTACGACA 655  
|||||  
Db 884 AGAATTTGAAGTCCGAAATATCATTTGCTCCATAAGGAATTCGATGATGACACTTACGACA 943  
|||||  
QY 656 ATGACATTCGCGTGTGTCAGCTGAAATCGGATTCGCTCCGCTGTGCCAGGAGACGCG 715  
Db 944 ATGACATTCGCGTGTGTCAGCTGAAATCGGATTCGCTCCGCTGTGCCAGGAGACGCG 1003  
|||||  
QY 716 TGGTCCGCACTGTGTGCTTCCCGCGGACCTGCGAGTCCGCACTGGAGCGAGTGTG 775  
Db 1004 TGGTCCGCACTGTGTGCTTCCCGCGGACCTGCGAGTCCGCACTGGAGCGAGTGTG 1063  
|||||  
QY 776 AGCTCTCCGGCTTACGCGCAAGCATGAGGCTTCTCTCTTCTTATTCGAGCGGCTGAAGG 835  
Db 1064 AGCTCTCCGGCTTACGCGCAAGCATGAGGCTTCTCTCTTCTTATTCGAGCGGCTGAAGG 1123  
|||||  
QY 836 AGGCTCATGTGACATGTACCCATCCAGCGCTGCACATCAACAATTTACTTAACAGAA 895  
Db 1124 AGGCTCATGTGACATGTACCCATCCAGCGCTGCACATCAACAATTTACTTAACAGAA 1183  
|||||  
QY 896 CAGTCACGACAAATGCTGTGTGCTGAGACACTCGAGCGCGGCGGCCCGCAGCAAACT 955  
Db 1184 CAGTCACGACAAATGCTGTGTGCTGAGACACTCGAGCGCGGCGGCCCGCAGCAAACT 1243  
|||||  
QY 956 TGCACGACCTGCCAGGCGGATTCGGAGGCGCCCTGCTGTGCTGAACGATGCGCGCA 1015  
Db 1244 TGCACGACCTGCCAGGCGGATTCGGAGGCGCCCTGCTGTGCTGAACGATGCGCGCA 1303  
|||||  
QY 1016 TGACTTTGGTGGGCAATCATAGCTGGGCGCTGGGCTGTGGACAGAGATGTCGCGGTG 1075  
Db 1304 TGACTTTGGTGGGCAATCATAGCTGGGCGCTGGGCTGTGGACAGAGATGTCGCGGTG 1363  
|||||  
QY 1076 TGACACAAAGTTACCAACTACCTAGACTGATGATGATGATGATGATGATGATGATG 1128  
Db 1364 TGACACAAAGTTACCAACTACCTAGACTGATGATGATGATGATGATGATGATGATG 1416  
|||||

RESULT 7  
AR059996 AR059996 1419 bp DNA linear PAT 29-SEP-1999  
LOCUS Sequence 62 from patent US 5840533.  
DEFINITION AR059996  
ACCESSION AR059996  
VERSION AR059996.1 GI:5986446  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE Unclassified.  
AUTHORS Niwa,M., Saito,Y., Sasaki,H., Hayashi,M., Notani,J. and Kobayashi,M.  
TITLE Tissue plasminogen activator  
JOURNAL Patent: US 5840533-A 62 24-NOV-1998;  
FEATURES Location/Qualifiers  
source 1..1419  
/organism="unknown"  
BASE COUNT 309 a 416 c 423 g 271 t  
ORIGIN

Query Match 94.4%; Score 1065; DB 6; Length 1419;  
Best Local Similarity 99.5%; Pred. No. 6.5e-244;  
Matches 1068; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
QY 56 CCCAGGCGGCTCTGAGGAAACAGTGTACTTTGGGAATGGGTTCAGCCCTACCGTG 115  
Db 344 CCCCTGCTGTCTGAGGAAACAGTGTACTTTGGGAATGGGTTCAGCCCTACCGTG 403  
QY 116 GCACGACAGCTCACCGAGTCGGTGTCTCTGCTCCCGTGGAAATCCCATGATCTCTGA 175  
Db 404 GCACGACAGCTCACCGAGTCGGTGTCTCTGCTCCCGTGGAAATCCCATGATCTCTGA 463  
QY 176 TAGGCAAGGTTTACACACAGACAGAACCCAGTGTGCCAGGCACTGGCGCTGGGCAACATA 235  
|||||

Db 464 TAGCAAGGTTTACACAGCACAGAACCCAGTGTCCAGGCACCTGGGCTGGGCAACATA 523  
QY 236 ATTAATCCCGGAATCCCTGATGGGATGCCAAGCCCTGGTCCACGTCGCTGAAGAACGCA 295  
Db 524 ATTAATCCCGGAATCCCTGATGGGATGCCAAGCCCTGGTCCACGTCGCTGAAGAACGCA 583  
|||||  
QY 296 GGCTGACGCTGGGAGTACTGTGATGTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 355  
Db 584 GGCTGACGCTGGGAGTACTGTGATGTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 643  
|||||  
QY 356 GCCAGCTCATGTTTCGATCAAAAGAGGGCTCTTTCGCCGACATCGCTCCACCCCTGGC 415  
Db 644 GCCAGCTCATGTTTCGATCAAAAGAGGGCTCTTTCGCCGACATCGCTCCACCCCTGGC 703  
|||||  
QY 416 AGGTCGCATCTTTTCCCAAGCACAGGAGGTGCGCCGAGAGCGGTTCTGTGCGGGGGCA 475  
Db 704 AGGTCGCATCTTTTCCCAAGCACAGGAGGTGCGCCGAGAGCGGTTCTGTGCGGGGGCA 763  
|||||  
QY 476 TACTCATCAGCTCTCTGCTGGATTCCTCTGCGCGCCCACTGCTTCCAGGAGAGGTTTCCGC 535  
Db 764 TACTCATCAGCTCTCTGCTGGATTCCTCTGCGCGCCCACTGCTTCCAGGAGAGGTTTCCGC 823  
|||||  
QY 536 CCACACACTGACGGTGTCTTGGGACAGAAACATACCGGGTGGTCTCTGGGAGAGAGC 595  
Db 824 CCACACACTGACGGTGTCTTGGGACAGAAACATACCGGGTGGTCTCTGGGAGAGAGC 883  
|||||  
QY 596 AGAATTTGAAGTCGAAATATCATTTGCTCCATAAGGAATTCGATGATGACACTTACGACA 655  
Db 884 AGAATTTGAAGTCGAAATATCATTTGCTCCATAAGGAATTCGATGATGACACTTACGACA 943  
|||||  
QY 656 ATGACATTCGCGTGTGCTGACCTGAAATCGGATTCGCTCCCGCTGTGCCAGGAGACGCG 715  
Db 944 ATGACATTCGCGTGTGCTGACCTGAAATCGGATTCGCTCCCGCTGTGCCAGGAGACGCG 1003  
|||||  
QY 716 TGGTCCGCACTGTGTGCTTCCCGCGGACCTGCGAGCTGCCGGAATGGAACGAGTGTG 775  
Db 1004 TGGTCCGCACTGTGTGCTTCCCGCGGACCTGCGAGCTGCCGGAATGGAACGAGTGTG 1063  
|||||  
QY 776 AGCTCTCCGGCTACGCGCAAGCATGAGGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 835  
Db 1064 AGCTCTCCGGCTACGCGCAAGCATGAGGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1123  
|||||  
QY 836 AGGCTCATGTGACATGTACCCATCCAGCGCTGCACATCAACAATTTACTTAACAGAA 895  
Db 1124 AGGCTCATGTGACATGTACCCATCCAGCGCTGCACATCAACAATTTACTTAACAGAA 1183  
|||||  
QY 896 CAGTCACGACAAATGCTGTGTGCTGAGACACTCGAGCGCGGCGGCCCGCAGCAAACT 955  
Db 1184 CAGTCACGACAAATGCTGTGTGCTGAGACACTCGAGCGCGGCGGCCCGCAGCAAACT 1243  
|||||  
QY 956 TGCACGACCTGCCAGGCGGATTCGGAGGCGCCCTGCTGTGCTGAACGATGCGCGCA 1015  
Db 1244 TGCACGACCTGCCAGGCGGATTCGGAGGCGCCCTGCTGTGCTGAACGATGCGCGCA 1303  
|||||  
QY 1244 TGACACAAAGTTACCAACTACCTAGACTGATGATGATGATGATGATGATGATGATG 1363  
Db 1364 TGACACAAAGTTACCAACTACCTAGACTGATGATGATGATGATGATGATGATGATG 1416  
|||||

RESULT 8  
E01944 E01944 1419 bp DNA linear PAT 29-SEP-1997  
LOCUS Synthetic DNA encoding new t-PA tissue plasminogen activator.  
DEFINITION E01944  
ACCESSION E01944  
VERSION E01944.1 GI:2170192  
KEYWORDS JP 1989104167-A/11.  
SOURCE synthetic construct  
ORGANISM synthetic sequences.  
artificial sequences.



REFERENCE	1 (bases 1 to 1419)	
AUTHORS	Niwa, M., Saico, Y., Sasaki, H., Hayashi, M., Notani, J. and Kobayashi, M.	
TITLE	NOVEL TISSUE PLASMINOGEN ACTIVATION FACTOR	
JOURNAL	Patent: JP 1989104167-A 11 21-APR-1989;	
COMMENT	FUJISAWA PHARMACEUT CO LTD	
OS	Artificial gene	
OC	Artificial sequence; Genes.	
PN	JP 1989104167-A/11	
PD	21-APR-1989	
PR	01-AUG-1988 JP 1988192320	
PR	03-AUG-1987 GB 87 8725052, PR	
PI	13-NOV-1987 GB 87 8726683	
PI	NIWA MINEO, SAITO YOSHIMASA, SASAKI HITOSHI, HAYASHI MASAKO, NOTANI JOJI,	
PI	KOBAYASHI MASAKAZU	
PC	C12N9/64, C12N1/20, C12N5/00, C12N15/00//A61K37/54, (C12N9/64, PC C12R1.19),	
PC	(C12N9/64, C12R1.91);	
CC	strandedness: Double;	
CC	topology: Linear;	
CC	hypothetical: No;	
CC	anti-sense: No;	
CC	*source: clone-pmTQk112;	
FH	Key	
FH	Location/Qualifiers	
FT	CDS	
FT	1..1419	
FT	Location/Qualifiers	
FT	/product='new t-PA'.	
FEATURES	1..1419	
source	/organism='synthetic construct'	
	/mol_type='genomic DNA'	
	/db_xref='taxon:32630'	
BASE COUNT	309 a 413 c 426 g 271 t	
ORIGIN		
Query Match	94.4%; Score 1065; DB 6; Length 1419;	
Best Local Similarity	99.5%; Pred. No. 6.5e-244;	
Matches 1068; Conservative	0; Mismatches 5; Indels 0; Gaps 0;	
QY	56 CCCAGCGCGCTCTGAGGAAACAGTACTGCTACTTTGGGAATGGGTTCAGCCTACCGTG 115	
DB	344 CCCCTGCTCTGAGGAAACAGTACTGCTACTTTGGGAATGGGTTCAGCCTACCGTG 403	
QY	116 GCACGACAGCTCACCAGAGTCGGTCTCTGCTCCCGTGGAAATCCATGATCTCTGA 175	
DB	404 GCACGACAGCTCACCAGAGTCGGTCTCTGCTCCCGTGGAAATCCATGATCTCTGA 463	
QY	176 TAGGCAAGTTTACACAGACAGAACCCAGTGCAGGCACTGGGCTGGGCAACATA 235	
DB	464 TAGGCAAGTTTACACAGACAGAACCCAGTGCAGGCACTGGGCTGGGCAACATA 523	
QY	236 ATTACTGCGGAATCTGATGGGATGCCAGCCCTGGTGCACGTGCTGAAGAACCGCA 295	
DB	524 ATTACTGCGGAATCTGATGGGATGCCAGCCCTGGTGCACGTGCTGAAGAACCGCA 583	
QY	296 GGCTGACGTGGGAGTACTGTGATGTCCTCTGCTCCACCTGGCGCTGAGACAGTACA 355	
DB	584 GGCTGACGTGGGAGTACTGTGATGTCCTCTGCTCCACCTGGCGCTGAGACAGTACA 643	
QY	356 GCCAGCCTCAGTTTCGATCAAGAGGGCTCTTCGCCGACATGCCCTCCACCCCTGGC 415	
DB	644 GCCAGCCTCAGTTTCGATCAAGAGGGCTCTTCGCCGACATGCCCTCCACCCCTGGC 703	
QY	416 AGGCTGCCATCTTTGCCAAGCAGAGAGTGCCTGGGAGAGGGTTCCTGTGCGGGGCA 475	
DB	704 AGGCTGCCATCTTTGCCAAGCAGAGAGTGCCTGGGAGAGGGTTCCTGTGCGGGGCA 763	
QY	476 TACTCATCAGCTCTCTGCTGGATTCCTCTGCTGGCGCCACTGCTTCAGGAGAGGTTTCGC 535	
DB	764 TACTCATCAGCTCTCTGCTGGATTCCTCTGCTGGCGCCACTGCTTCAGGAGAGGTTTCGC 823	
QY	536 CCCACCCTGACCGTGTATCTTTGGGCAGAACATACCCGGGTGGTCCCTGGCGAGGAGGAGC 595	
Db	824 CCCACCCTGACCGTGTATCTTTGGGCAGAACATACCCGGGTGGTCCCTGGCGAGGAGGAGC 883	
QY	596 AGAAATTTGAAGTCGAAAAATACATTTGTCCATAAGGAATTCGATGATGACACTTTACGACA 655	
DB	884 AGAAATTTGAAGTCGAAAAATACATTTGTCCATAAGGAATTCGATGATGACACTTTACGACA 943	
QY	656 ATGACATTTGGCTGCTGCAGCTGAAATCGAATTCGATTCCTCCGCTGTGCCAGAGAGCAGCG 715	
DB	944 ATGACATTTGGCTGCTGCAGCTGAAATCGAATTCGATTCCTCCGCTGTGCCAGAGAGCAGCG 1003	
QY	716 TGGTCCGCACTGTGTGCTTTCCCGCGGAGACCTGCGAGCTGCCGAGACTGACGAGGTGTG 775	
DB	1004 TGGTCCGCACTGTGTGCTTTCCCGCGGAGACCTGCGAGCTGCCGAGACTGACGAGGTGTG 1063	
QY	776 AGCTCTCCGGCTACGCAAGCATGAGCGCTTGTCTCTCTTCTATTTCGGAGCGGCTGAAGG 835	
DB	1064 AGCTCTCCGGCTACGCAAGCATGAGCGCTTGTCTCTCTTCTATTTCGGAGCGGCTGAAGG 1123	
QY	836 AGGCTCATGTGACACTGTATCCCATCCAGCCGCTGCACATCAACAATTTACTTAACAGAA 895	
DB	1124 AGGCTCATGTGACACTGTATCCCATCCAGCCGCTGCACATCAACAATTTACTTAACAGAA 1183	
QY	896 CAGTCACGCAACATCTGTGTGCTGAGACACTCGAGGCGGCGGCCCCAGGCAAACT 955	
DB	1184 CAGTCACGCAACATCTGTGTGCTGAGACACTCGAGGCGGCGGCCCCAGGCAAACT 1243	
QY	956 TGACACGAGCGCTGCAGGCGGATTTCCGGAGGCGCCCTGGTGTGTCTCAACGATGGCGCA 1015	
DB	1244 TGACACGAGCGCTGCAGGCGGATTTCCGGAGGCGCCCTGGTGTGTCTCAACGATGGCGCA 1303	
QY	1016 TGACTTTTGGTGGGCATCATCAGCTGGGGCCCTGGGCTGTGGACAGAGGATGTCCCGGTG 1075	
DB	1304 TGACTTTTGGTGGGCATCATCAGCTGGGGCCCTGGGCTGTGGACAGAGGATGTCCCGGTG 1363	
QY	1076 TGACACAAAGGTTTACCAACTACTAGACTGGATTCGTGACAACTGCGACCG 1128	
DB	1364 TGACACAAAGGTTTACCAACTACTAGACTGGATTCGTGACAACTGCGACCG 1416	
RESULT 9		
LOCUS	I06609 1689 bp DNA linear PAT 02-DEC-1994	
DEFINITION	Sequence 44 from Patent EP 0293934.	
ACCESSION	I06609	
VERSION	I06609.1 GI:590639	
KEYWORDS	Unknown.	
SOURCE	Unknown.	
ORGANISM	Unclassified.	
REFERENCE	1 (bases 1 to 1689)	
AUTHORS	Mulvihill, E.R., Nexo, B.A., Yoshitake, S., Ikeda, Y., Suzuki, S., Hashimoto, A. and Yuzuriha, T.	
TITLE	Mutant t-PA with kringle replacement	
JOURNAL	Patent: EP 0293934-A1 44 07-DEC-1988;	
FEATURES	Location/Qualifiers	
source	1..1689	
	/organism='unknown'	
BASE COUNT	377 a 483 c 505 g 324 t	
ORIGIN		
Query Match	94.4%; Score 1065; DB 6; Length 1689;	
Best Local Similarity	99.5%; Pred. No. 6.5e-244;	
Matches 1068; Conservative	0; Mismatches 5; Indels 0; Gaps 0;	
QY	56 CCCAGCGCGCTCTGAGGAAACAGTACTGCTACTTTGGGAATGGGTTCAGCCTACCGTG 115	
DB	614 CCCCTGCTCTGAGGAAACAGTACTGCTACTTTGGGAATGGGTTCAGCCTACCGTG 673	
QY	116 GCACGACAGCTCACCAGAGTCCGGTCCCTCTGCTCCCGTGGAAATTCATGATCTCTGA 175	
DB	674 GCACGACAGCTCACCAGAGTCCGGTCCCTCTGCTCCCGTGGAAATTCATGATCTCTGA 733	

Qy	176	TAGC	AAGT	TTT	TAC	CAG	CA	GCA	CG	CC	AG	TG	CC	CA	GG	CA	CT	TG	GG	CC	TG	GG	CA	AA	CA	T	235	
Db	734	TAGC	AA	AGT	TTT	TAC	AG	CA	CG	CC	AG	TG	CC	CA	GG	CA	CT	TG	GG	CC	TG	GG	CA	AA	CA	T	793	
Qy	236	ATT	ACT	CC	CG	AAT	CT	CAT	TG	GG	GAT	CC	CA	AG	CC	CT	TG	TC	CA	CG	TG	TC	CA	AG	AA	C	295	
Db	794	ATT	ACT	CC	CG	AAT	CT	CAT	TG	GG	GAT	CC	CA	AG	CC	CT	TG	TC	CA	CG	TG	TC	CA	AG	AA	C	853	
Qy	296	GG	CT	AG	CT	GG	GA	T	CT	GT	GA	TG	CC	CT	CT	CT	TC	CA	CT	GG	CC	T	G	AG	CA	G	355	
Db	854	GG	CT	AG	CT	GG	GA	T	CT	GT	GA	TG	CC	CT	CT	CT	TC	CA	CT	GG	CC	T	G	AG	CA	G	913	
Qy	356	GC	AG	CC	CT	CA	GT	TT	CG	CA	T	CA	AA	AG	AG	GG	CT	CT	TG	CC	GA	CA	TG	CC	CA	C	415	
Db	914	GC	AG	CC	CT	CA	GT	TT	CG	CA	T	CA	AA	AG	AG	GG	CT	CT	TG	CC	GA	CA	TG	CC	CA	C	973	
Qy	416	AG	GT	CC	CA	CT	TT	TG	CC	AA	GC	AC	AG	AG	GT	CG	CC	GA	AG	GG	CT	CT	TG	CC	GA	G	475	
Db	974	AG	GT	CC	CA	CT	TT	TG	CC	AA	GC	AC	AG	AG	GT	CG	CC	GA	AG	GG	CT	CT	TG	CC	GA	G	1033	
Qy	476	T	A	C	T	A	T	C	A	G	T	C	C	T	G	T	G	A	T	T	C	T	T	C	T	C	T	1034
Db	1034	T	A	C	T	A	T	C	A	G	T	C	C	T	G	T	G	A	T	T	C	T	T	C	T	C	T	1093
Qy	536	C	C	A	C	A	C	T	G	A	C	T	T	G	G	C	A	A	C	A	T	A	C	G	G	T	G	595
Db	1094	C	C	A	C	A	C	T	G	A	C	T	T	G	G	C	A	A										

[illegible]

QY 896 CAGTACCCGACAACTGCTGTGTGTGAGACACTCGAGCGCGCGGCCCGGCAAACT 955  
Db 1454 CAGTACCCGACAACTGCTGTGTGTGAGACACTCGAGCGCGCGGCCCGGCAAACT 1513  
QY 956 TGCACGACGCTCCAGGCGGATTCGGAGGCGCCCTGTGTGTGAGACAGATGCGCGCA 1015  
Db 1514 TGCACGACGCTCCAGGCGGATTCGGAGGCGCCCTGTGTGTGAGACAGATGCGCGCA 1573  
QY 1016 TGACTTTGGTGGGCAATCATACAGTGGGCGCTGGGTGTGAGACAGATGCTCCGGGTG 1075  
Db 1574 TGACTTTGGTGGGCAATCATACAGTGGGCGCTGGGTGTGAGACAGATGCTCCGGGTG 1633  
QY 1076 TGACACAAAGTTACCACTACCTAGACTGATTCGAGCAAACTGCGACG 1128  
Db 1634 TGACACAAAGTTACCACTACCTAGACTGATTCGAGCAAACTGCGACG 1686

RESULT 11  
108789  
LOCUS 108789 1689 bp DNA linear PAT 02-DEC-1994  
DEFINITION Sequence 3 from Patent WO 8804690.  
ACCESSION 108789  
VERSION 108789.1 GI:588494  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 1689)  
AUTHORS Bollen,A.J., Gheysen,D., Jacobs,P., Pierard,L. and Collen,D.J.  
JOURNAL Patent: WO 8804690-A 3 30-JUN-1988;  
FEATURES Location/Qualifiers  
source 1..1689  
/organism="unknown"  
BASE COUNT 379 a 480 c 505 g 325 t  
ORIGIN

Query Match 94.4%; Score 1065; DB 6; Length 1689;  
Best Local Similarity 99.5%; Pred. No. 6.5e-244;  
Matches 1068; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 56 CCCAGCGCGCTCTCAGGAAACAGTACTGCTTTGGGAATGGGTACGCTACCGTG 115  
Db 614 CCCCTGCTCTCTGAGGAAACAGTACTGCTTTGGGAATGGGTACGCTACCGTG 673  
QY 116 GCACGACAGCTCTACCGAGTGGGTGCTCTGCTCCCGTGGAAATCCCATGATCTGA 175  
Db 674 GCACGACAGCTCTACCGAGTGGGTGCTCTGCTCCCGTGGAAATCCCATGATCTGA 733  
QY 176 TAGGCAAGTTTACACAGACAGAACCCGAGTCCCGAGGCACTGGGCTGGGCAACATA 235  
Db 734 TAGGCAAGTTTACACAGACAGAACCCGAGTCCCGAGGCACTGGGCTGGGCAACATA 793  
QY 236 ATTACTGCGGGAATCCTGATGGGATGCAGAGCTGCTGCTGCTGAGACGCA 295  
Db 794 ATTACTGCGGGAATCCTGATGGGATGCAGAGCTGCTGCTGCTGAGACGCA 853  
QY 296 GCTGACGTGGGAGTACTGATGTCCTCTGCTCCACCTGCGGCTGAGACAGTACA 355  
Db 854 GCTGACGTGGGAGTACTGATGTCCTCTGCTCCACCTGCGGCTGAGACAGTACA 913  
QY 356 GCCAGCTCAGTTTCCGATCAAGAGAGGCTTCTGCGCGACATGCTCCCGCTGGC 415  
Db 914 GCCAGCTCAGTTTCCGATCAAGAGAGGCTTCTGCGCGACATGCTCCCGCTGGC 973  
QY 416 AGGCTGCCATCTTCCGACGACAGAGTGGCGCGGAGGCTTCTGCTGCGGCGCA 475  
Db 974 AGGCTGCCATCTTCCGACGACAGAGTGGCGCGGAGGCTTCTGCTGCGGCGCA 1033  
QY 476 TACTCATCAGTCTCTGCTGATCTCTCTGCGCGCACTGCTTCCAGGAGGTTTCCGC 535  
Db 1034 TACTCATCAGTCTCTGCTGATCTCTCTGCGCGCACTGCTTCCAGGAGGTTTCCGC 1093  
QY 536 CCCACCACTGACGGTATCTTTGGGCAACATACCGGGTGGTCCCTGGCGAGGAGGAGC 595

Db 1094 CCCACCACTGACGGTGTCTTTGGGCAACATACCGGGTGGTCCCTGGCGAGGAGGAGC 1153  
QY 596 AGAAATTTGAAGTTCGAAATAATACATTTGTCATAAGGAATTCGATGATGACACTTACGACA 655  
Db 1154 AGAAATTTGAAGTTCGAAATAATACATTTGTCATAAGGAATTCGATGATGACACTTACGACA 1213  
QY 656 ATGACATTTGCGCTCTGCGAGCTGAAATCGGATTCGTCCTGCTGCTCCCGAGGAGGAGCAG 715  
Db 1214 ATGACATTTGCGCTCTGCGAGCTGAAATCGGATTCGTCCTGCTGCTCCCGAGGAGGAGCAG 1273  
QY 716 TGGTCCGACACTGTGTGCTTCCCGCGGAGCTGCGAGCTGCGGAGCTGCGAGCTGCGAGTGTG 775  
Db 1274 TGGTCCGACACTGTGTGCTTCCCGCGGAGCTGCGAGCTGCGGAGCTGCGAGTGTG 1333  
QY 776 AGCTCTCGGCTACGGCAAGCATGAGGCTTGTCTCTCTTTCTATTTCGAGCGGCTGAAGG 835  
Db 1334 AGCTCTCGGCTACGGCAAGCATGAGGCTTGTCTCTCTTTCTATTTCGAGCGGCTGAAGG 1393  
QY 836 AGGCTCATGTGACACTGTATCCCATCCAGCGCTGCACATCAACAATTTACTTTAACAGAA 895  
Db 1394 AGGCTCATGTGACACTGTATCCCATCCAGCGCTGCACATCAACAATTTACTTTAACAGAA 1453  
QY 896 CAGTACCGACAACTGCTGTGTGCTGAGACACTCGAGCGCGCGGCCCGGCAAACT 955  
Db 1454 CAGTACCGACAACTGCTGTGTGCTGAGACACTCGAGCGCGCGGCCCGGCAAACT 1513  
QY 956 TGCACGACGCTCTCAGGCGGATTCGGAGGCGCCCTGCTGTGTCTGAACCATGCGCGCA 1015  
Db 1514 TGCACGACGCTCTCAGGCGGATTCGGAGGCGCCCTGCTGTGTCTGAACCATGCGCGCA 1573  
QY 1016 TGACTTTGGTGGGCACTCATCAGTGGGCGCTGGGCTGTGAGACAGAGGATGCTCCGGGTG 1075  
Db 1574 TGACTTTGGTGGGCACTCATCAGTGGGCGCTGGGCTGTGAGACAGAGGATGCTCCGGGTG 1633  
QY 1076 TGTACACAAAGTTTACCACTACCTAGACTGATTCGAGCAAACTGCGACG 1128  
Db 1634 TGTACACAAAGTTTACCACTACCTAGACTGATTCGAGCAAACTGCGACG 1686

RESULT 12  
A06611  
LOCUS A06611 1780 bp DNA linear PAT 23-JUN-1993  
DEFINITION Synthetic nucleotide sequence of the 1.78kb t-PA BamHI fragment of pt-PABam.  
ACCESSION A06611  
VERSION A06611.1 GI:412906  
KEYWORDS  
SOURCE synthetic construct  
ORGANISM synthetic construct  
REFERENCE 1 (bases 1 to 1780)  
AUTHORS  
JOURNAL  
FEATURES Location/Qualifiers  
source 1..1780  
/organism="synthetic construct"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:32630"  
17..1705  
/codon\_start=1  
/product="t-PA analogue"  
/protein\_id="CAA00577.1"  
/db\_xref="GI:412907"

Patent: WO 900600-A 11 25-JAN-1990;  
1..1780  
Location/Qualifiers  
source

CDS  
17..1705  
/translation="MDAMKRLCCVLLCGAVFVPSQEIHARPRGARSQVTCRDE  
KTFMIYQHQSWLRFVLRNRYEWCNRSRAQCHSVFKSCSPFCNFGTCCQALY  
YDFVCPGEGFAGKCEIDTRATCYEDOGISYRGWTSTAESGAECTNWNSSALAOXP  
YSGRRPDAIRLGLNHNCRNPDSDKPCYVFKAGYSSEFCSTPACSGNSDCYFG  
NSAYRGTHSLTESGASCLPWNSSMLIGKYTAQNPQAQALGLKHNCRNPDGAPR  
WCHVKNRRTWEYCDVPSCTGLRQYSQFRIKGLGFADTASHPWQAIAFGRHR  
SPGRFLCGGILLISSCWILSAHCFQERFPFHLTVILGRTYRVVPGEEQKEVEKY  
IVHKEFDDDDTDNDIALQLKSDSSRCAQESSVVRVTVCLPPADLQLPDWTECELSGYG

RHEALSPFYSERLKEAHRVLYPSSRCTSOHLNRTVTDNMLCAGDTRSGGPGQANLHDA  
CQDGGGGLVCLNDGRMTLVGLIISWGLGCKQDVPVYTKVNYLDWIRDNRP"  
misc\_difference 371  
/note="t" can be replaced by 'a'.  
misc\_difference 1367.1369  
/note="aag" can be replaced by 'agt' or 'tcc'.  
misc\_difference 1370.1372  
/note="cat" can be replaced by 'act'.  
BASE COUNT 405 a 513 c 523 g 339 t  
ORIGIN

Query Match 94.4%; Score 1065; DB 6; Length 1780;  
Best Local Similarity 99.5%; Pred. No. 6.5e-244;  
Matches 1068; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
QY 56 CCCAGCGGCTCTGAGGGAACAGTGTACTTTGGGATGGGTGAGCTACCGTG 115  
DB 630 CCCCTGCTCTGAGGGAACAGTGTACTTTGGGATGGGTGAGCTACCGTG 689  
QY 116 GCACGACAGCTCACCGAGTCGGTCTCTGCTCCCGTGGAAATTCATGATCTGA 175  
DB 690 GCACGACAGCTCACCGAGTCGGTCTCTGCTCCCGTGGAAATTCATGATCTGA 749  
QY 176 TAGGCAAGTTTACACGACAGAACCCAGTGCCTGAGGCACTGGGCTGGGCAACATA 235  
DB 750 TAGGCAAGTTTACACGACAGAACCCAGTGCCTGAGGCACTGGGCTGGGCAACATA 809  
QY 236 ATTACTCGCGGATCTGATGGGATCCAGCCCTGTGTCAGCTGTGAAGACCGCA 295  
DB 810 ATTACTCGCGGATCTGATGGGATCCAGCCCTGTGTCAGCTGTGAAGACCGCA 869  
QY 296 GGCTGACGTGGAGTACTGTGATGCTCTGCTCCACCTGCGGCTGAGACAGTACA 355  
DB 870 GGCTGACGTGGAGTACTGTGATGCTCTGCTCCACCTGCGGCTGAGACAGTACA 929  
QY 356 GCCAGCTCAGTTTCGATCAAGAGAGGCTTTCGCGGCACTGCTCCACCCCTGGC 415  
DB 930 GCCAGCTCAGTTTCGATCAAGAGAGGCTTTCGCGGCACTGCTCCACCCCTGGC 989  
QY 416 AGGCTGCCATCTTCCGACGACAGAGTGCCTGAGAGCGGTCTGTCGGGGGCA 475  
DB 990 AGGCTGCCATCTTCCGACGACAGAGTGCCTGAGAGCGGTCTGTCGGGGGCA 1049  
QY 476 TACTCATCAGCTCTGCTGGATTCTCTGCGGCGCACTGCTTCCAGGAGGTTTCGC 535  
DB 1050 TACTCATCAGCTCTGCTGGATTCTCTGCGGCGCACTGCTTCCAGGAGGTTTCGC 1109  
QY 536 CCCACACCTGACGTGATCTTTGGGCAAGACATACCGGTGTCCTCGGAGGAGGAGC 595  
DB 1110 CCCACACCTGACGTGATCTTTGGGCAAGACATACCGGTGTCCTCGGAGGAGGAGC 1169  
QY 596 AGAAATTTGAAGTCGAAATATCATTTGCTTCCATAGGAATTCGATGATGACACTTACGACA 655  
DB 1170 AGAAATTTGAAGTCGAAATATCATTTGCTTCCATAGGAATTCGATGATGACACTTACGACA 1229  
QY 656 ATGACATTCGCTGCTGACGTGAAATCGGATTCGCTCCGCTGTCGCCAGGAGACGCG 715  
DB 1230 ATGACATTCGCTGCTGACGTGAAATCGGATTCGCTCCGCTGTCGCCAGGAGACGCG 1289  
QY 716 TGGTCCGACATGTGTGCTTCCCGGCGGACCTGACGTCCGACCTGGAAGGAGTG 775  
DB 1290 TGGTCCGACATGTGTGCTTCCCGGCGGACCTGACGTCCGACCTGGAAGGAGTG 1349  
QY 776 AGCTCTCGGCTACGGCAGCATGAGCTGTGCTCTTCTTATTCGAGCGGCTGAAGG 835  
DB 1350 AGCTCTCGGCTACGGCAGCATGAGCTGTGCTCTTCTTATTCGAGCGGCTGAAGG 1409  
QY 836 AGGCTCATGTGACATGTACCATCCAGCGCTGACATCAACATTTACTTAACAGAA 895  
DB 1410 AGGCTCATGTGACATGTACCATCCAGCGCTGACATCAACATTTACTTAACAGAA 1469  
QY 896 CAGTCAACGACAAATGTGTGTGTGAGACACTTCGAGGCGGCGGCCCGAGCAAACT 955

Db 1470 CAGTCAACGACAAATGTGTGTGTGTGAGACACTCGGAGCGGCGGCCAGCAAACT 1529  
QY 956 TGCACGACGCTGCCAGGCGATTTCGGAGGCGCCCTGTGTGTCTGAACGATGCGCGCA 1015  
Db 1530 TGCACGACGCTGCCAGGCGATTTCGGAGGCGCCCTGTGTGTCTGAACGATGCGCGCA 1589  
QY 1016 TGACTTTGGTGGCATCATCAGCTGGGCGCTGGGCTGTGGACAGAAGATGTCGGGTG 1075  
Db 1590 TGACTTTGGTGGCATCATCAGCTGGGCGCTGGGCTGTGGACAGAAGATGTCGGGTG 1649  
QY 1076 TGACACAAAGTTTACCAACTACCTAGACTGGATTCGTGACAAATGCGGACG 1128  
Db 1650 TGACACAAAGTTTACCAACTACCTAGACTGGATTCGTGACAAATGCGGACG 1702

RESULT 13  
LOCUS A10226 1780 bp DNA linear PAT 17-FEB-1997  
DEFINITION Synthetic DNA for mutant tissue plasminogen activator (t-PA).  
ACCESSION A10226  
VERSION A10226.1 GI:490657  
KEYWORDS synthetic construct  
SOURCE synthetic construct  
ORGANISM artificial sequences.  
REFERENCE 1 (bases 1 to 1780)  
AUTHORS Petersen, L.C. and Boel, E.  
TITLE A tissue plasminogen activator analogue  
JOURNAL Patent: EP 0351246-A 7 17-JAN-1990;  
NOVO-NORDISK A/S; NOVO NORDISK A/S  
FEATURES  
source  
1. .1780  
/organism="synthetic construct"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:32630"  
17. .1705  
/codon\_start=1  
/transl\_table=11  
/product="mutant tissue plasminogen activator (t-PA)"  
/protein\_id="CAA00883.1"  
/db\_xref="GI:490658"  
/translation="MDAMKRGCLCCVLLCGAVFVSPSQEIHARFRGARSQVTCRDE  
KTMIVYQHQSWLRPVLRNRYCWNRSRQACHSVFVKSCSPRCFNGGTCCQALY  
FSDFCPCPGFAGKCEIDTRATCYEDQISYRGTWSTAEGAECTNMSSALAKP  
YSGRRPDAIRLGNHNYCRNDRSKPCVFKAGVYSEFCSTPACSENGSDCYFG  
NGSAYRGTSLTESGASCLPWSMILLIKVYTAQNPQAALGKHNKCRNPDGAKP  
WCHVKNRRTWEICDVPSCSTGLROYSQPRFKGLFADIAHPWQAIAFAHRR  
SPGERFJCGGILISSCWILSAHCFQERFPFHLTVILGRTRYVVPGESEKQFVEKY  
IVHKEFDDTDNDIALQLKSSRCAQESSVTVTVLPADLPQDPTWTECELSYG  
KHEALSPFYSERLKEAHRVLYPSSRCTSOHLNRTVTDNMLCAGDTRSGGPGQANLHDA  
CQDGGGGLVCLNDGRMTLVGLIISWGLGCKQDVPVYTKVNYLDWIRDNRP"

CDS  
misc\_difference 371.373  
/note="in t-PA C87S, K419S"  
/replace="agt"  
misc\_difference 1367.1372  
/note="in t-PA K419S, H420T"  
/replace="agtact"  
misc\_difference 1367.1369  
/note="in t-PA C87S, K419S"  
/replace="tcc"  
BASE COUNT 405 a 513 c 523 g 339 t  
ORIGIN

Query Match 94.4%; Score 1065; DB 6; Length 1780;  
Best Local Similarity 99.5%; Pred. No. 6.5e-244;  
Matches 1068; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
QY 56 CCCAGGCGGCTCTGAGGGAACAGTGTACTTTGGGATGGGTGAGCTACCGTG 115  
DB 630 CCCCTGCTCTGAGGGAACAGTGTACTTTGGGATGGGTGAGCTACCGTG 689  
QY 116 GCACGACAGCTCACCGAGTCGGTCTCTGCTCCCGTGGAAATTCATGATCTGA 175

Db 690 GCACGCACAGCCTCACCAGTCCGGTGCCTCTCGCTCCGCTGGGAATTCATGATCCTGA 749  
Qy 176 TAGGCAAGTTTACACAGACAGAACCCAGAGTCCAGGCACTGGGCTGGGCAACATA 235  
Db 750 TAGGCAAGTTTACACAGACAGAACCCAGTCCAGGCACTGGGCTGGGCAACATA 809  
Qy 236 ATTACTGCGGATCCTGATGGGATGCAAGCCCTGTGTGCACTGCTGCAAGAACCGCA 295  
Db 810 ATTACTGCGGATCCTGATGGGATGCAAGCCCTGTGTGCACTGCTGCAAGAACCGCA 869  
Qy 296 GGTGACGTGGGAGTACTGTGATGTCCTCTCTGCTCCACCTGCGGCTGGAGACAGTACA 355  
Db 870 GGTGACGTGGGAGTACTGTGATGTCCTCTCTGCTCCACCTGCGGCTGGAGACAGTACA 929  
Qy 356 GCCAGCCTCAGTTTCGATCAAGAGGCTCTTCCGCGACATCGCTCCACCCCTGGC 415  
Db 930 GCCAGCCTCAGTTTCGATCAAGAGGCTCTTCCGCGACATCGCTCCACCCCTGGC 989  
Qy 416 AGGCTGCCATCTTTCGCAAGCACAGAGGTGCGCCGGAGAGCGTTCCTGTGCGGGGCA 475  
Db 990 AGGCTGCCATCTTTCGCAAGCACAGAGGTGCGCCGGAGAGCGTTCCTGTGCGGGGCA 1049  
Qy 476 TACTCATCAGCTCCTGCTGATCTCTCTGCGCCCACTGCTTCCAGGAGAGTTTCCGC 535  
Db 1050 TACTCATCAGCTCCTGCTGATCTCTCTGCGCCCACTGCTTCCAGGAGAGTTTCCGC 1109  
Qy 536 CCCACCACTGACGCTGATCTTGGGCAAGACATACCGGCTGCTTCCGCGAGGAGGAGC 595  
Db 1110 CCCACCACTGACGCTGATCTTGGGCAAGACATACCGGCTGCTTCCGCGAGGAGGAGC 1169  
Qy 596 AGAATTTGAGTCGAAATACATGTCCTATAGGAATTCATGATGACACTTACGACA 655  
Db 1170 AGAATTTGAGTCGAAATACATGTCCTATAGGAATTCATGATGACACTTACGACA 1229  
Qy 656 ATGACATTCGCTGCTGACGTGAATCGATTCCTGCGGCTGCTTCCGCGAGGAGGAGC 715  
Db 1230 ATGACATTCGCTGCTGACGTGAATCGATTCCTGCGGCTGCTTCCGCGAGGAGGAGC 1289  
Qy 716 TGGTCCGACATGTGTGCTTCCCGCGGACCTGACGCTGCGGACCTGGAGCGAGTGTG 775  
Db 1290 TGGTCCGACATGTGTGCTTCCCGCGGACCTGACGCTGCGGACCTGGAGCGAGTGTG 1349  
Qy 776 AGCTCTCGGCTACGGCAAGCATGAGGCTTGTCTCTTCTTATTCGGAGCGGCTGAAGG 835  
Db 1350 AGCTCTCGGCTACGGCAAGCATGAGGCTTGTCTCTTCTTATTCGGAGCGGCTGAAGG 1409  
Qy 836 AGGCTCATGTGACAGCTGACCATCCAGCCGCTGCACATCAACATTTACTTACAGAA 895  
Db 1410 AGGCTCATGTGACAGCTGACCATCCAGCCGCTGCACATCAACATTTACTTACAGAA 1469  
Qy 896 CAGTCACCGACAACATGTGTGTGTGGAGACACTCGGAGCGGCGGCGCCAGGCAAACT 955  
Db 1470 CAGTCACCGACAACATGTGTGTGTGGAGACACTCGGAGCGGCGGCGCCAGGCAAACT 1529  
Qy 956 TGCAGACGCTGCCAGGCGAATTCGGAGGCGGCGGCTGTGTGTGAAACGATGCGCGCA 1015  
Db 1530 TGCAGACGCTGCCAGGCGAATTCGGAGGCGGCGGCTGTGTGTGAAACGATGCGCGCA 1589  
Qy 1016 TGACTTTGTGGGCATCATCAGCTGGGCGCTGGGCTGTGGACAGAGGATGTCCCGGTG 1075  
Db 1590 TGACTTTGTGGGCATCATCAGCTGGGCGCTGGGCTGTGGACAGAGGATGTCCCGGTG 1649  
Qy 1076 TGTACAAAAGTTTACCAACTACCTAGACTGGATTGCTGCAACATGCGACCG 1128  
Db 1650 TGTACAAAAGTTTACCAACTACCTAGACTGGATTGCTGCAACATGCGACCG 1702.

RESULT 14  
E01176  
LOCUS E01176 1836 bp DNA linear PAT 29-SEP-1997  
DEFINITION DNA encoding tris-kringle plasminogen activator.  
ACCESSION E01176  
VERSION E01176.1 GI:2169435

KEYWORDS JP 1987104577-A/1.  
SOURCE synthetic construct  
ORGANISM synthetic construct  
REFERENCE 1 (bases 1 to 1836)  
AUTHORS Pooru,P.H., Narendaa,K.K. and Shiyouguangu,R.R.  
TITLE POLYCLONAL PLASMINOGEN ACTIVATING FACTOR  
JOURNAL Patent: JP 1987104577-A 1 15-MAY-1987;  
AMERICAN HOME PROD CORP  
COMMENT OS Artificial gene  
OC Artificial sequence; Genes.  
OS homo sapiens  
PN JP 1987104577-A/1  
PD 15-MAY-1987  
PF 13-AUG-1986 JP 1986191173  
PR 14-AUG-1985 US 85 766163, 11-JUL-1986 US 86 884835 PI  
POORU POOVEN HANGU, NARENDAA KUMAARU KARUVAN, PI SHIYOUGUANGU RIN  
R11 C12N9/52,A61K35/12,A61K35/74,A61K37/02,A61K37/54,C07H21/04, PC  
C12N15/00,  
PC (C12N9/52,C12R1:19),(C12N15/00,C12R1:19);  
CC strandedness: Double;  
CC topology: Linear;  
CC hypothetical: No;  
CC anti-sense: No;  
CC \*source: cell\_type=fibroblast;  
CC \*source: cell\_line=WI-38cell;  
FH Key Location/Qualifiers  
FT CDS 12..1793  
FT sig\_peptide 12..71  
FT mat\_peptide 72..1790  
FT /product='this-kringle plasminogen activator'.  
FEATURES Location/Qualifiers  
source 1..1836  
/organism="synthetic construct"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:32630"  
BASE COUNT 422 a 537 c 528 g 349 t  
ORIGIN  
Query Match 94.4%; Score 1065; DB 6; Length 1836;  
Best Local Similarity 99.5%; Pred. No. 6.5e-244;  
Matches 1068; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
Qy 56 CCCAGGCGGCTCTGAGGGAACAGTCACTGCTACTTTGGGAATGGGTACGCTACCGTG 115  
Db 718 CCCCTGCTGCTCTGAGGGAACAGTCACTGCTACTTTGGGAATGGGTACGCTACCGTG 777  
Qy 116 GCAGGCACAGCCTCACCAGTCCGGTGCCTCTGCTCCCTCCGCTGGAATTCATGATCCTGA 175  
Db 778 GCAGGCACAGCCTCACCAGTCCGGTGCCTCTGCTCCCTCCGCTGGAATTCATGATCCTGA 837  
Qy 176 TAGGCAAGTTTACACAGCACAGAACCCAGTGCACAGGCACTGGGCTGGGCAACATA 235  
Db 838 TAGGCAAGTTTACACAGCACAGAACCCAGTGCACAGGCACTGGGCTGGGCAACATA 897  
Qy 236 ATTTACTCGGATTCCTGATGGGATGCCAAGCCCTGGTGCACGCTGCTGAGAACCGCA 295  
Db 898 ATTTACTCGGATTCCTGATGGGATGCCAAGCCCTGGTGCACGCTGCTGAGAACCGCA 957  
Qy 296 GGCTGAGCTGGGAGTACTGTGATGTCCTGCTCCACTCGGCTGGAGACAGTACA 355  
Db 958 GGCTGAGCTGGGAGTACTGTGATGTCCTGCTCCACTCGGCTGGAGACAGTACA 1017  
Qy 356 GCCAGCCTCAGTTTCGATCAAGAGGAGGCTCTTTCGCGGACATCGCTCCACCCCTGGC 415  
Db 1018 GCCAGCCTCAGTTTCGATCAAGAGGAGGCTCTTTCGCGGACATCGCTCCACCCCTGGC 1077  
Qy 416 AGGCTGCCATCTTTCGCAAGCACAGAGGTGCGCCGGAGAGCGTTCCTGTGCGGGGCA 475

Db 1078 AGGCTGCCACTCTTTGCCAAGCACAGAGAGTCCGCCGAGAGCGGTTCTGTGCGGGGCA 1137  
Qy 476 TACTCATCAGCTCCCTGCTGGATTCTCTCTGCGGCCCACTGCTTCCAGAGAGGTTTCGCG 535  
Db 1138 TACTCATCAGCTCCCTGCTGGATTCTCTCTGCGGCCCACTGCTTCCAGAGAGGTTTCGCG 1197  
Qy 536 CCCACCACCTGACCGGTGATCTTTGGCGAGAACATACCCGGGTGCTCCTGCGGAGGAGGAGC 595  
Db 1198 CCCACCACCTGACCGGTGATCTTTGGCGAGAACATACCCGGGTGCTCCTGCGGAGGAGGAGC 1257  
Qy 596 AGAAATTTGAAGTCGAAATATACATTTGTCATAGGAATTCGATGACACTTTACGACA 655  
Db 1258 AGAAATTTGAAGTCGAAATATACATTTGTCATAGGAATTCGATGACACTTTACGACA 1317  
Qy 656 ATGACATTTGCGCTGCTGACGCTGAAATCGGATTCGTCCTGCTGCGGAGGAGGAGC 715  
Db 1318 ATGACATTTGCGCTGCTGACGCTGAAATCGGATTCGTCCTGCTGCGGAGGAGGAGC 1377  
Qy 716 TGGTCCGCACTGTGTGCTTCCCGGGGACCTGACGCTGCGGACCTGGACGCGAGTGTG 775  
Db 1378 TGGTCCGCACTGTGTGCTTCCCGGGGACCTGACGCTGCGGACCTGGACGCGAGTGTG 1437  
Qy 776 AGCTCTCGGCTGACGCGAAGATGAGGCTTGTCTCTCTTCTATTCGAGGCGGCTGAAGG 835  
Db 1438 AGCTCTCGGCTGACGCGAAGATGAGGCTTGTCTCTCTTCTATTCGAGGCGGCTGAAGG 1497  
Qy 836 AGGCTCATGTGACACTGTACCCATCCAGCGCTGCACATCAACATTTACTTAACAGAA 895  
Db 1498 AGGCTCATGTGACACTGTACCCATCCAGCGCTGCACATCAACATTTACTTAACAGAA 1557  
Qy 896 CAGTCACCGCAAACTGTGTGTGCTGGAGACACTCGGAGCGCGGGCCCGAGGCAAACT 955  
Db 1558 CAGTCACCGCAAACTGTGTGTGCTGGAGACACTCGGAGCGCGGGCCCGAGGCAAACT 1617  
Qy 956 TGCACGACGCTGCCAGGCGGATTCGGAGGCGCCCTGCTGCTGCTGAACGATGCGCGCA 1015  
Db 1618 TGCACGACGCTGCCAGGCGGATTCGGAGGCGCCCTGCTGCTGCTGAACGATGCGCGCA 1677  
Qy 1016 TGACTTTGCTGGGCATCATCAGCTGGCGCTGGGCTGTGGACAGAGGATGTCGCGGTG 1075  
Db 1678 TGACTTTGCTGGGCATCATCAGCTGGCGCTGGGCTGTGGACAGAGGATGTCGCGGTG 1737  
Qy 1076 TGTACACAAAGGTTACCACTACCTAGACTGGAATTCGTCGACAACTGCGAGCG 1128  
Db 1738 TGTACACAAAGGTTACCACTACCTAGACTGGAATTCGTCGACAACTGCGAGCG 1790

## RESULT 15

LOCUS 101581 1836 bp ss-DNA linear PAT 21-MAY-1993  
DEFINITION Sequence 1 from Patent US 4916071.  
ACCESSION 101581

VERSION 101581.1 GI:271057

KEYWORDS

SOURCE

ORGANISM

Unknown.

Unclassified.

REFERENCE 1 (bases 1 to 1836)

AUTHORS Hung, P.P., Kalyan, N.K. and Lee, S.-g.I.

TITLE Poly-kringle plasminogen activator

JOURNAL Patent: US 4916071-A 1 10-APR-1990;

American Home Products Corporation; New York, NY

FEATURES

source 1..1836

/organism="unknown"

BASE COUNT 422 a 537 c 528 g 349 t

ORIGIN

Query Match 94.4%; Score 1065; DB 6; Length 1836;

Best Local Similarity 99.5%; Pred. No. 6.5e-244;

Matches 1068; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 56 CCCAGGCGGCTCTGAGGGAACAGTGTACTTGTGGGAATGGGTTCAGCCTACCGTG 115

Db 718 CCCCTGCTGCTCTGTAGGGAAAACAGTGACTGTCTACTTTTGGGAATGGGTTCAGCCTACCGTG 777  
Qy 116 GCAGCGACAGCTCACCAGAGTCGGGTGCTCTCTGCTCCCGTGGAAATTCATGATCTCTGA 175  
Db 778 GCAGCGACAGCTCACCAGAGTCGGGTGCTCTCTGCTCCCGTGGAAATTCATGATCTCTGA 837  
Qy 176 TAGGCAAGGTTTACACAGCACAGAAACCCAGTGCAGGACCTGCGGCTGGGCAAAACATA 235  
Db 838 TAGGCAAGGTTTACACAGCACAGAAACCCAGTGCAGGACCTGCGGCTGGGCAAAACATA 897  
Qy 236 ATTAATCTCCGGAATTCCTGATGGGATGCCAAGCCCTGGTGCACGTGCTGAAGAACCGCA 295  
Db 898 ATTAATCTCCGGAATTCCTGATGGGATGCCAAGCCCTGGTGCACGTGCTGAAGAACCGCA 957  
Qy 296 GGCTGACGTGGGAGTACTGTGATGTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 355  
Db 958 GGCTGACGTGGGAGTACTGTGATGTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1017  
Qy 356 GCCAGCCTCAGTTCGATCAAAAGGAGGGCTCTTTCCGCGACATCGCTCCACCCCTGGC 415  
Db 1018 GCCAGCCTCAGTTCGATCAAAAGGAGGGCTCTTTCCGCGACATCGCTCCACCCCTGGC 1077  
Qy 416 AGGCTGCCATCTTTTGCACAGCACAGGAGTGCCTCCGCGAGAGCGGTTCTGTGCGGGGCA 475  
Db 1078 AGGCTGCCATCTTTTGCACAGCACAGGAGTGCCTCCGCGAGAGCGGTTCTGTGCGGGGCA 1137  
Qy 476 TACTCATCAGCTCTCTGCTGATTTCTCTGCGCGCCACTGCTTCCAGAGAGGTTTCGCG 535  
Db 1138 TACTCATCAGCTCTCTGCTGATTTCTCTGCGCGCCACTGCTTCCAGAGAGGTTTCGCG 1197  
Qy 536 CCACACCTCAGCGTGCATCTTGGGAGAAACATACCGGGTGGTCCCTGGCGAGGAGGAGC 595  
Db 1198 CCACACCTCAGCGTGCATCTTGGGAGAAACATACCGGGTGGTCCCTGGCGAGGAGGAGC 1257  
Qy 596 AGAAATTTGAAGTCGAAATATACATTTGTCATAGGAATTCGATGATGACACTTTACGACA 655  
Db 1258 AGAAATTTGAAGTCGAAATATACATTTGTCATAGGAATTCGATGATGACACTTTACGACA 1317  
Qy 656 ATGACATTTGCGCTGCTGACGCTGAAATCGGATTCGTCCTCTTCTATTCGAGGCGGCTGA 715  
Db 1318 ATGACATTTGCGCTGCTGACGCTGAAATCGGATTCGTCCTCTCTCTCTCTCTCTCTCTCT 1377  
Qy 716 TGGTCCGCACTGTGTGCTTCCCGGGGACCTGACGCTGCGGACCTGGACGCGAGTGTG 775  
Db 1378 TGGTCCGCACTGTGTGCTTCCCGGGGACCTGACGCTGCGGACCTGGACGCGAGTGTG 1437  
Qy 776 AGCTCTCCGCTACGCGAAGCATGAGGCTTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 835  
Db 1438 AGCTCTCCGCTACGCGAAGCATGAGGCTTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1497  
Qy 836 AGGCTCATGTGACACTGTACCCATCCAGCGCTGCACATCAACATTTACTTAACAGAA 895  
Db 1498 AGGCTCATGTGACACTGTACCCATCCAGCGCTGCACATCAACATTTACTTAACAGAA 1557  
Qy 896 CAGTCACCGCAAACTGTGTGTGCTGGAGACACTCGGAGCGCGGGCCCGAGGCAAACT 955  
Db 1558 CAGTCACCGCAAACTGTGTGTGCTGGAGACACTCGGAGCGCGGGCCCGAGGCAAACT 1617  
Qy 956 TGCACGACGCTGCCAGGCGGATTCGGAGGCGCCCTGCTGCTGCTGAACGATGCGCGCA 1015  
Db 1618 TGCACGACGCTGCCAGGCGGATTCGGAGGCGCCCTGCTGCTGCTGAACGATGCGCGCA 1677  
Qy 1016 TGACTTTGCTGGGCATCATCAGCTGGCGCTGGGCTGTGGACAGAGGATGTCGCGGTG 1075  
Db 1678 TGACTTTGCTGGGCATCATCAGCTGGCGCTGGGCTGTGGACAGAGGATGTCGCGGTG 1737  
Qy 1076 TGTACACAAAGGTTACCACTACCTAGACTGGAATTCGTCGACAACTGCGAGCG 835  
Db 1738 TGTACACAAAGGTTACCACTACCTAGACTGGAATTCGTCGACAACTGCGAGCG 1497

Search completed: August 19, 2003, 04:46:34

Job time : 4737.31 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 18, 2003, 19:51:53 ; Search time 397.386 Seconds  
(without alignments)  
7662.485 Million cell updates/sec

Title: US-09-987-455-5  
Perfect score: 1128  
Sequence: 1 atgaaaagacagtatgcg.....ttcgtgacaacatcgaccg 1128

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_19Jun03.\*  
1: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.\*  
2: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.\*  
3: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.\*  
4: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.\*  
5: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.\*  
6: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.\*  
7: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.\*  
8: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.\*  
9: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.\*  
10: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.\*  
11: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.\*  
12: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.\*  
13: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.\*  
14: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.\*  
15: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.\*  
16: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.\*  
17: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.\*  
18: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.\*  
19: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.\*  
20: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.\*  
21: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.\*  
22: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.\*  
23: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.\*  
24: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.\*  
25: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1128	100.0	1128	24	AAD40614
2	1065	94.4	1314	10	AA991123
3	1065	94.4	1341	11	AAQ06762
4	1065	94.4	1419	10	AA991133
5	1065	94.4	1689	9	AA982178
6	1065	94.4	1780	11	AAQ01358
7	1065	94.4	2100	10	AA991119
8	1065	94.4	2100	11	AAQ05532

9	1065	94.4	2162	16	AAQ87370	Human tPA clone.
10	1065	94.4	2162	19	AAV37294	Human tissue plasminogen activator; tPA; kringle 2 serine protease; stroke;
11	1065	94.4	2547	7	AA960659	Sequence encoding
12	1065	94.4	2560	10	AA990542	Plasminogen
13	1065	94.4	7533	8	AA970491	Entire sequence of
14	1065	94.4	7533	9	AA981970	Plasminogen
15	1064	94.3	1680	17	AA972588	Novel plasminogen
16	1064	94.3	1680	17	AA972587	Novel plasminogen
17	1063.4	94.3	1614	11	AAQ03581	Glycosylated plasminogen
18	1063.4	94.3	1614	12	AAQ11550	Sequence encoding
19	1063.4	94.3	1659	12	AAQ11551	Sequence encoding
20	1063.4	94.3	1686	21	ABL52893	Tissue plasminogen
21	1063.4	94.3	1689	9	AA982179	Mutated t-PA analog
22	1063.4	94.3	1689	9	AA981090	Part of tPA024 gen
23	1063.4	94.3	1689	11	AAQ04903	Part of tPA023 gen
24	1063.4	94.3	1689	11	AAQ04904	Sequence encoding
25	1063.4	94.3	1689	11	AAQ05177	T-PA67+ mutant wit
26	1063.4	94.3	1689	12	AAQ13379	Human tissue plasminogen
27	1063.4	94.3	1689	24	ABK52546	Full-length tissue plasminogen
28	1063.4	94.3	1740	17	AA972585	Human tissue plasminogen
29	1063.4	94.3	1955	21	AA923198	Sequence encoding
30	1063.4	94.3	1989	7	AA960700	T-PA variant contg
31	1063.4	94.3	2091	12	AAQ12073	Plasminogen
32	1063.4	94.3	2099	11	AAQ05534	T-PA with -ve char
33	1063.4	94.3	2100	12	AAQ12072	Sequence encoding
34	1063.4	94.3	2170	16	AA970248	Human tissue plasminogen
35	1063.4	94.3	2170	16	AAQ86576	Plasminogen
36	1063.4	94.3	2249	11	AAQ05535	Sequence encoding
37	1063.4	94.3	2313	11	AAQ05944	cDNA sequence enco
38	1063.4	94.3	2457	7	AA960146	cDNA of human tiss
39	1063.4	94.3	2509	10	AA990508	Human tissue plasminogen
40	1063.4	94.3	2519	24	ABK52545	Human tissue type
41	1063.4	94.3	2519	24	ABL59544	Human tissue type
42	1063.4	94.3	2547	4	AA930001	Sequence encoding
43	1063.4	94.3	7360	17	AA915930	DHFR/intron (WTras
44	1062	94.1	1065	24	AAD40616	Human kringle 2 se
45	1062	94.1	1068	10	AA991120	Sequence of coding

ALIGNMENTS

RESULT 1  
AAD40614  
ID AAD40614 standard; DNA; 1128 BP.

AC AAD40614;

DT 30-OCT-2002 (first entry)

XX OmpA-K2S fusion protein encoding DNA.

XX Tissue plasminogen activator; tPA; kringle 2 serine protease; stroke;  
XX K2S; cardiac infarction; acute myocardial infarction; pulmonary embolism;  
XX artery occlusion; deep vein thrombosis; blood clotting; thrombolytic;  
XX cerebroprotective; cardiac; ompA; fusion protein; gene; ds.

OS Unidentified.

XX Key Location/Qualifiers

FT CDS 1..1128

FT FT /\*tag= a

FT FT /product= "OmpA-K2S fusion protein"

FT FT /transl\_except= (pos:1126..1128, aa:Pro-Gly)

FT FT /note= "No stop codon"

FT FT /partial

XX W0200240650-A2.

XX 23-MAY-2002.

XX 07-NOV-2001; 2001WO-EPI2857.

PR 14-NOV-2000; 2000GB-0027779.  
XX (BOEH ) BOEHRINGER INGELHEIM INT GMBH.  
XX Werner R, Goetz F, Tayapiwatana C, Manosroi J, Manosroi A;  
PI WPI; 2002-519376/55.  
DR P-PSDB; AAE25034.  
XX  
XX Producing active, correctly folded recombinant tissue plasminogen  
PT activator, kringle 2 serine protease in prokaryotic cells by expressing  
PT the protein-encoding DNA operably linked to DNA coding for signal  
PT peptide OmpA -  
XX  
XX Claim 7; Page 30; 80pp; English.  
XX  
XX The present invention relates to a method of producing extracellularly  
CC secreted, active, correctly folded, recombinant tissue plasminogen  
CC activator (tPA), kringle 2 serine protease molecule (K2S) or their  
CC variants in prokaryotic cells by expressing the protein-encoding DNA  
CC operably linked to DNA coding for signal peptide OmpA. The method is  
CC useful for producing recombinant DNA-derived tPA, K2S or their variants.  
CC Sequences of the invention are useful for manufacturing a medicament  
CC for treating stroke, cardiac infarction, acute myocardial infarction,  
CC pulmonary embolism, any artery occlusion such as intracranial artery  
CC occlusion (e.g., arteries supplying the brain), peripherally occluded  
CC arteries, coronary artery occlusion, deep vein thrombosis or related  
CC diseases associated with unwanted blood clotting. The present sequence  
CC is a DNA encoding a fusion protein comprising OmpA and K2S protein.  
XX  
XX Sequence 1128 BP; 238 A; 332 C; 332 G; 226 T; 0 other;  
Query Match 100.0%; Score 1128; DB 24; Length 1128;  
Best Local Similarity 100.0%; Pred. No. 5.1e-239;  
Matches 1128; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ATGAAAAGACAGCTATCGGATTCAGTGGCAGCTGGCTGTTTCGCTACCGTGGCCAG 60  
DB 1 ATGAAAAGACAGCTATCGGATTCAGTGGCAGCTGGCTGTTTCGCTACCGTGGCCAG 60  
QY 61 GCGGCTCTGAGGAAACAGTGAATGCTTTTGGGAATGGGTACGCTACCGTGGCAG 120  
DB 61 GCGGCTCTGAGGAAACAGTGAATGCTTTTGGGAATGGGTACGCTACCGTGGCAG 120  
QY 121 CACAGCTCAGCGAGTGGGTGCTCTCCCTCCGTCGGAATTCATGATCCTGATAGGC 180  
DB 121 CACAGCTCAGCGAGTGGGTGCTCTCCCTCCGTCGGAATTCATGATCCTGATAGGC 180  
QY 181 AAGGTTTACACAGCAGAAACCCAGTCCAGGCACTGGGCTGGGCAACATAATTAC 240  
DB 181 AAGGTTTACACAGCAGAAACCCAGTCCAGGCACTGGGCTGGGCAACATAATTAC 240  
QY 241 TCCCGGAATCTGATGGGATGCCAAGCCCTGGTGCCACGTCGTGAAGAACCCGAGGCTG 300  
DB 241 TCCCGGAATCTGATGGGATGCCAAGCCCTGGTGCCACGTCGTGAAGAACCCGAGGCTG 300  
QY 301 ACGTGGGAGTACTGTGATGTCCTCCCTGCTCCACCTGGGCTGAGACATAGCAGCAG 360  
DB 301 ACGTGGGAGTACTGTGATGTCCTCCCTGCTCCACCTGGGCTGAGACATAGCAGCAG 360  
QY 361 CCTCAGTTTCGATCAAAAGGAGGGCTCTTCGCGCAGACATCGCTCCACCCCTGGCAGCT 420  
DB 361 CCTCAGTTTCGATCAAAAGGAGGGCTCTTCGCGCAGACATCGCTCCACCCCTGGCAGCT 420  
QY 421 GCCATCTTTGCCAAGCAGAGAGTGGCCCGGAGAGCGGTTTCCTGTGCGGGGCACTACTC 480  
DB 421 GCCATCTTTGCCAAGCAGAGAGTGGCCCGGAGAGCGGTTTCCTGTGCGGGGCACTACTC 480  
QY 481 ATCAGCTCCTGCTGATTTCTCTGCGCGCCACATGCTTCCAGGAGAGGTTTCCGCCCCAC 540  
DB 481 ATCAGCTCCTGCTGATTTCTCTGCGCGCCCCACATGCTTCCAGGAGAGGTTTCCGCCCCAC 540  
QY 541 CACCTGACGCTGATCTTGGGAGAAACATACCGGGTGGTCCCTGGCGAGGAGGAGAGAA 600

DB 541 CACCTGACGCTGATCTTGGGAGAAACATACCGGGTGGTCCCTGGCGAGGAGGAGAGAA 600  
QY 601 TTTGAAGTCGAAATAATACATTGCTTCCATGAAGAAATTCAGATGATGACACTTACGACAAATGAC 660  
DB 601 TTTGAAGTCGAAATAATACATTGCTTCCATGAAGAAATTCAGATGATGACACTTACGACAAATGAC 660  
QY 661 ATTGCGCTGTCAGCTGAAATCGGATTCGTCCTGCTGCTCCCGAGGAGGAGGAGCGTGGTC 720  
DB 661 ATTGCGCTGTCAGCTGAAATCGGATTCGTCCTGCTGCTCCCGAGGAGGAGGAGCGTGGTC 720  
QY 721 CGCACTGTCGCTTCCCGCGGAGCTCGCAGCTGCGGACATGCGAGTGTGAGCTC 780  
DB 721 CGCACTGTCGCTTCCCGCGGAGCTCGCAGCTGCGGACATGCGAGTGTGAGCTC 780  
QY 781 TCCGGCTACGGCAAGCATGAGGCTTGTCTCTTTTATTTCGAGCGGCTGAAGAGGCT 840  
DB 781 TCCGGCTACGGCAAGCATGAGGCTTGTCTCTTTTATTTCGAGCGGCTGAAGAGGCT 840  
QY 841 CATGTGAGCTGTACCATCCAGCCGCTGCACATCAACATTTACTTAAACAGAACAGTC 900  
DB 841 CATGTGAGCTGTACCATCCAGCCGCTGCACATCAACATTTACTTAAACAGAACAGTC 900  
QY 901 ACCGACAACTGCTGTGCTGCTGAGACACTCGGAGCGCGGCGCCAGGCAAACTTGCAC 960  
DB 901 ACCGACAACTGCTGTGCTGCTGAGACACTCGGAGCGCGGCGCCAGGCAAACTTGCAC 960  
QY 961 GAGCGCTCCAGGCGGAGTTCGGAGGCGCCCTGCTGCTGAACGATGCGCGGCTGACT 1020  
DB 961 GAGCGCTCCAGGCGGAGTTCGGAGGCGCCCTGCTGCTGAACGATGCGCGGCTGACT 1020  
QY 1021 TTGGTGGGATCATCAGCTGGGCGCTGGGCTGTGGACAGAGATGTCGCGGCTGTGTAC 1080  
DB 1021 TTGGTGGGATCATCAGCTGGGCGCTGGGCTGTGGACAGAGATGTCGCGGCTGTGTAC 1080  
QY 1081 ACAAGGTTTACCAACTACCTAGACTGGATTTCGTGACAACTGCGACCG 1128  
DB 1081 ACAAGGTTTACCAACTACCTAGACTGGATTTCGTGACAACTGCGACCG 1128  
RESULT 2  
AAN91123  
ID AAN91123 standard; DNA; 1314 BP.  
XX AC AAN91123;  
XX 25-MAR-2003 (updated)  
DT 03-OCT-2002 (updated)  
DT 18-JUN-1990 (first entry)  
XX  
DE Sequence of coding region in plasmid pTQkPA delta trp.  
XX  
KW Tissue plasminogen activator; tPA; thrombolytic agent;  
KW plasminogen; vascular diseases.  
XX  
OS Synthetic.  
XX  
FH Key Location/Qualifiers  
FT CDS 1..1065  
FT /\*tag= a  
XX  
EP302456-A.  
PN  
XX  
PD 08-FEB-1989.  
XX  
XX 02-AUG-1988; 88EP-0112569.  
XX  
PR 03-AUG-1987; 87GB-0018298.  
PR 26-OCT-1987; 87GB-0025052.  
PR 13-NOV-1987; 87GB-0026683.  
XX  
PA (FUJI ) FUJISAWA PHARM CO LTD.  
XX

Db	1019	AGGCTCATGTGAGACTGTACCCATCCAGCGCGTGCATCAACATTTACTTTAACGAA	1078			
Qy	896	CAGTCACCGACAACATCTGTGTCTGGAGACACTCGGAGCGCGGCCCCAGGCAAACT	955			
Db	1079	CAGTCACCGACAACATCTGTGTCTGGAGACACTCGGAGCGCGGCCCCAGGCAAACT	1138			
Qy	956	TGCACGACGCTGCGAGGCGGATTCGGAGGCGGCCCTGTGTGTCTGAAACGATGCCGCA	1015			
Db	1139	TGCACGACGCTGCGAGGCGGATTCGGAGGCGGCCCTGTGTGTCTGAAACGATGCCGCA	1198			
Qy	1016	TGACTTTGTGGGATCATCAGCTGGGGCTTGGGCTGTGGACAGAGATGTCCCGGGTG	1075			
Db	1199	TGACTTTGTGGGATCATCAGCTGGGGCTTGGGCTGTGGACAGAGATGTCCCGGGTG	1258			
Qy	1076	TGTACACAAAGGTTACCAACTTACCTAGACTGGATTCTGTGACCAACATCGACCG	1128			
Db	1259	TGTACACAAAGGTTACCAACTTACCTAGACTGGATTCTGTGACCAACATCGACCG	1311			
RESULT 3						
AAQ06762						
ID	AAQ06762 standard; DNA; 1341 BP.					
XX	AAQ06762;					
XX	AC					
XX	XX					
DT	09-JAN-2003 (updated)					
DT	01-MAR-1991 (first entry)					
XX	Sequence encoding non-glycosylated tPA deriv.					
DE	XX					
XX	Tissue plasminogen activator; thrombolysis; finger domain;					
KW	EGF domain; ss.					
KW	XX					
XX	Synthetic.					
XX	XX					
FH	Key Location/Qualifiers					
FT	CDS 4..1341					
FT	/*tag= a					
FT	/product=tPA deriv.					
XX	EP400545-A.					
FN	XX					
XX	05-DEC-1990.					
PD	XX					
XX	28-MAY-1990; 90EP-0110096.					
PF	XX					
XX	14-JUL-1989; 89DE-3923339.					
PR	XX					
PR	31-MAY-1989; 89DE-3917781.					
XX	(BOEF ) BOEHRINGER MANNHEIM GMBH.					
PA	XX					
XX	Stern A, Kohnert U, Rudolph R, Fischer S, Martin U;					
PI	XX					
XX	WPI; 1990-363094/49.					
DR	P-PSDB; AAR08150.					
DR	XX					
XX	New non-glycosylated form of tissue plasminogen activator - with					
FT	thrombolytic activity and long plasma life					
FT	XX					
PS	Claim 2; page 16; 21pp; German.					
PS	XX					
XX	This sequence encodes a non-glycosylated tPA deriv. lacking the					
CC	finger- and EGF-domains. The polypeptide has a lower clearance					
CC	rate (longer half-life) than natural tPA while retaining thrombo-					
CC	lytic activity and stimulation by fibrin.					
CC	(Updated on 09-JAN-2003 to add missing OS field.)					
CC	XX					
XX	Sequence 1341 BP; 293 A; 397 C; 399 G; 252 T; 0 other;					
SQ	XX					
Query Match 94.4%; Score 1065; DB 11; Length 1341;						
Best Local Similarity 99.5%; Pred. No. 3.9e-225;						
Matches 1068; Conservative 0; Mismatches 5; Indels 0; Gaps						

```
QY 56 CCCAGCGCGCTCTGAGGAAACAGTACTGTCTACTTTGGGAATGGGTGAGCTACCGTG 115
Db 269 CCCCTGCTCTGAGGAAACAGTACTGTCTACTTTGGGAATGGGTGAGCTACCGTG 328
QY 116 GCACGACAGCTCTACCGAGTGGGTGCTCTGCTCCCGTGGAAATTCATGATCCTGA 175
Db 329 GCACGACAGCTCTACCGAGTGGGTGCTCTGCTCCCGTGGAAATTCATGATCCTGA 388
QY 176 TAGGCAAGTTTACACGACAGAACCCAGTGGCCAGGCACTGGCCCTGGGCAACATA 235
Db 389 TAGGCAAGTTTACACGACAGAACCCAGTGGCCAGGCACTGGCCCTGGGCAACATA 448
QY 236 ATTACTGCGGGAATCTGATGGGATGCCAAGCCCTGGTGCCACGTCTGAAGAACCGCA 295
Db 449 ATTACTGCGGGAATCTGATGGGATGCCAAGCCCTGGTGCCACGTCTGAAGAACCGCA 508
QY 296 GGCTGAGTGGGAGTACTGTGATGTGCTCTGCTCCACCTGGCGGCTGAGACAGTACA 355
Db 509 GGCTGAGTGGGAGTACTGTGATGTGCTCTGCTCCACCTGGCGGCTGAGACAGTACA 568
QY 356 GCCAGCCTCAGTTTCGATCAAGAGGGCTCTTCGCCGACATCGCCTCCACCCCTGGC 415
Db 569 GCCAGCCTCAGTTTCGATCAAGAGGGCTCTTCGCCGACATCGCCTCCACCCCTGGC 628
QY 416 AGGCTGCCATCTTTGCCAAGCACAGAGGTGGCCCGGAGAGCGGTTCCTGTGCGGGGCA 475
Db 629 AGGCTGCCATCTTTGCCAAGCACAGAGGTGGCCCGGAGAGCGGTTCCTGTGCGGGGCA 688
QY 476 TACTCATCAGCTCTGCTGATTTCTCTGCGGCGCACTGCTTCAGAGAGGTTTCGCG 535
Db 689 TACTCATCAGCTCTGCTGATTTCTCTGCGGCGCACTGCTTCAGAGAGGTTTCGCG 748
QY 536 CCCACACCTGACGTGATCTTGGGAGAACATACCGGGTGGTCCCTGGCAGGAGAGC 595
Db 749 CCCACACCTGACGTGATCTTGGGAGAACATACCGGGTGGTCCCTGGCAGGAGAGC 808
QY 596 AGAATTTGAAGTCGAAAAATACATTTGTCATAGGAATTCGATGATGACACTTACGACA 655
Db 809 AGAATTTGAAGTCGAAAAATACATTTGTCATAGGAATTCGATGATGACACTTACGACA 868
QY 656 ATGACATTCGCTGTGACGTGAAATCGGATTCGTCCTGCTGTCGCCAGAGAGCAGC 715
Db 869 ATGACATTCGCTGTGACGTGAAATCGGATTCGTCCTGCTGTCGCCAGAGAGCAGC 928
QY 716 TGGTCCGACATCTGTCCTCCCGGCGACCTGACGTCCGACCTGAGACGAGTGTG 775
Db 929 TGGTCCGACATCTGTCCTCCCGGCGACCTGACGTCCGACCTGAGACGAGTGTG 988
QY 776 AGCTCTCCGGCTACGGCAAGCATGAGGCTTTGTCTCTCTTTCTATTTCGGAGCGGTGAAG 835
Db 989 AGCTCTCCGGCTACGGCAAGCATGAGGCTTTGTCTCTCTTTCTATTTCGGAGCGGTGAAG 1048
QY 836 AGGCTCATGTGACGTGATCCCATCCAGCGCTGACATCAACATTTACTTAACAGAA 895
Db 1049 AGGCTCATGTGACGTGATCCCATCCAGCGCTGACATCAACATTTACTTAACAGAA 1108
QY 896 CAGTCACCGACAAATGTGTGTCTGAGACACTCCGAGCGGGGGCCCCCAGGCAACT 955
Db 1109 CAGTCACCGACAAATGTGTGTCTGAGACACTCCGAGCGGGGGCCCCCAGGCAACT 1168
QY 956 TGCACGACCTGCGCAGGCGGATTCGGAGCGCCCTGCTGTGTGTCTGAACATGGCGCA 1015
Db 1169 TGCACGACCTGCGCAGGCGGATTCGGAGCGCCCTGCTGTGTGTCTGNACATGGCGCA 1228
QY 1016 TGACTTTGGTGGGATCATAGCTGGGCTGGGCTGTGGACAGAGANTCCCGGGTG 1075
Db 1229 TGACTTTGGTGGGATCATAGCTGGGCTGGGCTGTGGACAGAGANTCCCGGGTG 1288
QY 1076 TGTACAAAGGTTACCACTTACCTAGACTGGATTCGTTGACAACTGGACCG 1128
Db 1289 TGTACAAAGGTTACCACTTACCTAGACTGGATTCGTTGACAACTGGACCG 1341
```

```
RESULT 4
AAN91133
ID AAN91133 standard; DNA; 1419 BP.
XX
XX AAN91133;
AC
XX 25-MAR-2003 (updated)
DT 03-OCT-2002 (updated)
DT 18-JUN-1990 (first entry)
XX
XX Sequence of coding region in plasmid pmrQk112.
XX
XX Tissue plasminogen activator; tPA; thrombolytic agent;
KW plasminogen; vascular diseases.
XX
XX Synthetic.
OS
XX
XX Key Location/Qualifiers
FH 1..1065
CDS /*tag= a
XX
XX EP302456-A.
FN
XX 08-FEB-1989.
PD
XX 02-AUG-1988; 88EP-0112569.
PF
XX 03-AUG-1987; 87GB-0018298.
PR 26-OCT-1987; 87GB-0025052.
PR 13-NOV-1987; 87GB-0026683.
XX
XX (FUJI ) FUJISAWA PHARM CO LTD.
XX
XX Niwa M, Saito Y, Sasaki H, Hayashi M, Notani J, Kobayashi M;
PI WPI; 1989-040625/06.
XX P-PSDB; AAP94416.
DR
XX
XX New tissue plasminogen activator -
PT comprising finger and growth factor domains lacking tPA for
PT longer half-life and stronger thrombolytic activity.
XX
XX Disclosure; Page ?; 68pp; English.
XX
XX (Updated on 03-OCT-2002 to add missing OS field.)
CC (Updated on 25-MAR-2003 to correct PA field.)
XX
XX Sequence 1419 BP; 309 A; 413 C; 426 G; 271 T; 0 other;
SQ
Query Match 94.4%; Score 1065; DB 10; Length 1419;
Best Local Similarity 99.5%; Pred. No. 3.9e-225;
Matches 1068; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 56 CCCAGCGCGCTCTGAGGAAACAGTACTGTCTACTTTGGGAATGGGTGAGCTACCGTG 115
Db 344 CCCCTGCTCTGAGGAAACAGTACTGTCTACTTTGGGAATGGGTGAGCTACCGTG 403
QY 116 GCACGACAGCTCTACCGAGTGGGTGCTCTGCTCCCGTGGAAATTCATGATCCTGA 175
Db 404 GCACGACAGCTCTACCGAGTGGGTGCTCTGCTCCCGTGGAAATTCATGATCCTGA 463
QY 176 TAGGCAAGTTTACACGACAGAACCCAGTGGCCAGGCACTGGCCCTGGGCAACATA 235
Db 464 TAGGCAAGTTTACACGACAGAACCCAGTGGCCAGGCACTGGCCCTGGGCAACATA 523
QY 236 ATTACTGCGGGAATCTGATGGGATGCCAAGCCCTGGTGCCACGTCTGAAGAACCGCA 295
Db 524 ATTACTGCGGGAATCTGATGGGATGCCAAGCCCTGGTGCCACGTCTGAAGAACCGCA 583
QY 296 GGCTGAGTGGGAGTACTGTGATGTGCTCTGCTCCACCTGGCGGCTGAGACAGTACA 355
Db 584 GGCTGAGTGGGAGTACTGTGATGTGCTCTGCTCCACCTGGCGGCTGAGACAGTACA 643
```

QY 356 GCCAGCCTCAGTTTCGCATCAAGAGAGGCTCTTCCGCGACATCCCTCCCAACCCCTGGC 415  
DB 644 GCCAGCCTCAGTTTCGCATCAAGAGAGGCTCTTCCGCGACATCCCTCCCAACCCCTGGC 703  
QY 416 AGGCTGCCATCTTTCGCCAGCACAGAGGTGCGCCGAGAGCGGTTCCTGTGCGGGGCA 475  
DB 704 AGGCTGCCATCTTTCGCCAGCACAGAGGTGCGCCGAGAGCGGTTCCTGTGCGGGGCA 763  
QY 476 TACTCATCAGCTCCTGCTGGATTCTCTCTGCGCGCCCACTGCTTCCAGGAGAGGTTTCGCG 535  
DB 764 TACTCATCAGCTCCTGCTGGATTCTCTCTGCGCGCCCACTGCTTCCAGGAGAGGTTTCGCG 823  
QY 536 CCCACCACTGACGGTGATCTTTGGGCGAACAATACCGGGTGCTCCCTGGGAGGAGGAGC 595  
DB 824 CCCACCACTGACGGTGATCTTTGGGCGAACAATACCGGGTGCTCCCTGGGAGGAGGAGC 883  
QY 596 AGAATTTGAGTCCGAAATACATGTTCCATAGGAATTCGATGATGACACTTACGACA 655  
DB 884 AGAATTTGAGTCCGAAATACATGTTCCATAGGAATTCGATGATGACACTTACGACA 943  
QY 656 ATGACATTCGCTGCTGACGCTGAAATCGGATTCGCTCCGCTGTGCCAGGAGAGCAGCG 715  
DB 944 ATGACATTCGCTGCTGACGCTGAAATCGGATTCGCTCCGCTGTGCCAGGAGAGCAGCG 1003  
QY 716 TGGTCCGCACTGTGTGCTTCCCGCGGACCTGACGCTGCGGACTGGAAGGAGTGTG 775  
DB 1004 TGGTCCGCACTGTGTGCTTCCCGCGGACCTGACGCTGCGGACTGGAAGGAGTGTG 1063  
QY 776 AGCTCTCGGCTACGGCAGCATGAGGCTTGTCTCTTCTATTCGAGCGGCTGAAGG 835  
DB 1064 AGCTCTCGGCTACGGCAGCATGAGGCTTGTCTCTTCTATTCGAGCGGCTGAAGG 1123  
QY 836 AGGCTCATGTGACAGTGTACCCATCCAGCGCTGCACATCAACATTTACTTAACAGAA 895  
DB 1124 AGGCTCATGTGACAGTGTACCCATCCAGCGCTGCACATCAACATTTACTTAACAGAA 1183  
QY 896 CAGTCACCGCAACATGTGTGTGCTGGAGACATTCGAGCGCGGGGCCCGAGCAAACT 955  
DB 1184 CAGTCACCGCAACATGTGTGTGCTGGAGACATTCGAGCGCGGGGCCCGAGCAAACT 1243  
QY 956 TGCAGCAGCCTGCCAGGCGGATTCGGAGAGCGCCCTGCTGTGTGCTGAACATGCGCGCA 1015  
DB 1244 TGCAGCAGCCTGCCAGGCGGATTCGGAGAGCGCCCTGCTGTGTGCTGAACATGCGCGCA 1303  
QY 1016 TGACTTTGGTGGGCATCATCAGCTGGGCGCTGGGCTGTGGACAGAGGATGTCCCGGCTG 1075  
DB 1304 TGACTTTGGTGGGCATCATCAGCTGGGCGCTGGGCTGTGGACAGAGGATGTCCCGGCTG 1363  
QY 1076 TGTACAAAAGTTTACCACTACCTAGACTGGATTGCTGACAAATGCGACCG 1128  
DB 1364 TGTACAAAAGTTTACCACTACCTAGACTGGATTGCTGACAAATGCGACCG 1416

RESULT 5  
AAN82178  
ID AAN82178 standard; DNA; 1689 BP.  
XX AC AAN82178;  
XX DT 25-MAR-2003 (updated)  
XX DT 10-MAR-2003 (updated)  
XX DT 03-NOV-1990 (first entry)  
XX DE Tissue plasminogen activator with S-119 subetd for M.  
XX KW Tissue plasminogen activator; tPA; thrombosis; N-glycosylation; ss.  
XX OS Unidentified.  
XX FH Key  
XX FT CDS 1..1689  
XX FT /\*tag= a

/product=modified tPA  
JP63230083-A.  
26-SEP-1988.  
20-MAR-1987; 87JP-0064339.  
20-MAR-1987; 87JP-0064339.  
(EISA ) EISAI CO LTD.  
WPI; 1988-311961/44.  
P-ESDB; AAP82581.  
Modified tissue plasminogen activator - having glycine-183 and serine-186 residues susd. with serine and threonine.  
Disclosure; Page ?; 16pp; Japanese.  
One N-glycosylation site, i.e. NSS (117-119) is substituted with NSM and the N-glycosylation is removed.  
Plasmid encoding the modified tPA is 99-6300 and its transformant is E.coli RRI-Zem 99-6300 (FERM P-9127).  
This modified tPA, used to treat thrombosis, is of high quality and has a longer half life period in blood.  
See also AAN82177-N82179.  
(Updated on 10-MAR-2003 to add missing OS field.)  
(Updated on 25-MAR-2003 to correct PF field.)  
(Updated on 25-MAR-2003 to correct PR field.)  
(Updated on 25-MAR-2003 to correct PA field.)  
XX SQ Sequence 1689 BP; 376 A; 482 C; 504 G; 327 T; 0 other;  
Query Match 94.4%; Score 1065; DB 9; Length 1689;  
Best Local Similarity 99.5%; Pred. No. 4e-225;  
Matches 1068; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
QY 56 CCCAGGCGGCTCTGAGGAAACAGTCACTGCTACTTTGGGAATGGGTGACGCTACCGTG 115  
DB 614 CCCCTGCTGCTCTGAGGAAACAGTCACTGCTACTTTGGGAATGGGTGACGCTACCGTG 673  
QY 116 GCACGACAGCCTCACCGAGTCGGGTGCTCTGCTCCGCTCCGATGAAATCCATGATCTGA 175  
DB 674 GCACGACAGCCTCACCGAGTCGGGTGCTCTGCTCCGCTCCGATGAAATCCATGATCTGA 733  
QY 176 TAGCAGAGTTTACACAGCACAGAACCCAGTCCGAGGACATGGGCTTGGGCAACATA 235  
DB 734 TAGCAGAGTTTACACAGCACAGAACCCAGTCCGAGGACATGGGCTTGGGCAACATA 793  
QY 236 ATTAATCCCGGAATCCCTGATGGGATGCCAAGCCCTGCTGCGACGCTGGAAGAACCGCA 295  
DB 794 ATTAATCCCGGAATCCCTGATGGGATGCCAAGCCCTGCTGCGACGCTGGAAGAACCGCA 853  
QY 296 GGCTGACGTGGGAGTACTGTGATGTGCCCTCTCTGCTCCACCTGCGGCTGAGACAGTACA 355  
DB 854 GGCTGACGTGGGAGTACTGTGATGTGCCCTCTCTGCTCCACCTGCGGCTGAGACAGTACA 913  
QY 356 GCCAGCCTCAGTTTCGCATCAAGAGAGGCTCTTTCGCGCGACATGCGCTCCCAACCCCTGGC 415  
DB 914 GCCAGCCTCAGTTTCGCATCAAGAGAGGCTCTTTCGCGCGACATGCGCTCCCAACCCCTGGC 973  
QY 416 AGGCTGCCATCTTTCGCCAAGCACAGGAGGTGCGCCGAGAGCGGTTCCTGTGCGGGGCA 475  
DB 974 AGGCTGCCATCTTTCGCCAAGCACAGGAGGTGCGCCGAGAGCGGTTCCTGTGCGGGGCA 1033  
QY 476 TACTCATCAGCTCCTGCTGGATTCTCTCTGCGCGCCCACTGCTTCCAGGAGAGGTTTCGCG 535  
DB 1034 TACTCATCAGCTCCTGCTGGATTCTCTCTGCGCGCCCACTGCTTCCAGGAGAGGTTTCGCG 1093  
QY 536 CCCACCACTGACGGTGATCTTTGGGCGAACAATACCGGGTGCTCCCTGGGAGGAGGAGC 595  
DB 1094 CCCACCACTGACGGTGATCTTTGGGCGAACAATACCGGGTGCTCCCTGGGAGGAGGAGC 1153

QY 596 AGAAATTTGAAGTCGAAATATACATTGTCATAGGAATTCGATGACACTTACGACA 655  
 DB 1154 AGAAATTTGAAGTCGAAATATACATTGTCATAGGAATTCGATGACACTTACGACA 1213  
 QY 656 ATGACATTTGCGCTGCTGAGCTGAAATCGGATTCGTCGCTGTCGCGGAGAGAGCG 715  
 DB 1214 ATGACATTTGCGCTGCTGAGCTGAAATCGGATTCGTCGCTGTCGCGGAGAGAGCG 1273  
 QY 716 TGGTCCGCACTGTGTGCTTCCCGCGGAGCTGCGAGCTGCGGAGAGAGAGTG 775  
 DB 1274 TGGTCCGCACTGTGTGCTTCCCGCGGAGCTGCGAGCTGCGGAGAGAGAGTG 1333  
 QY 776 AGCTCTCGGCTACGCGAAGATGAGGCTTGTCTCTTCTTATTCGAGGCGGTGAAG 835  
 DB 1334 AGCTCTCGGCTACGCGAAGATGAGGCTTGTCTCTTCTTATTCGAGGCGGTGAAG 1393  
 QY 836 AGGCTCATGTGAGCTGATACCATCCAGCGCTGCACATCAACATTTACTTAACAGAA 895  
 DB 1394 AGGCTCATGTGAGCTGATACCATCCAGCGCTGCACATCAACATTTACTTAACAGAA 1453  
 QY 896 CAGTCACGCAACATGCTGTGTGCTGAGACACTCGGAGGCGGCGGCGGCGGCAAACT 955  
 DB 1454 CAGTCACGCAACATGCTGTGTGCTGAGACACTCGGAGGCGGCGGCGGCGGCAAACT 1513  
 QY 956 TGCACGAGCGCTGCCAGGCGGATTCGGGAGGCGGCGGCTGTGTCTGAACGATGCGCGCA 1015  
 DB 1514 TGCACGAGCGCTGCCAGGCGGATTCGGGAGGCGGCGGCTGTGTCTGAACGATGCGCGCA 1573  
 QY 1016 TGACTTTGGTGGGCAATCATGCTGAGCTGGGCGCTGGGCTGTGGACAGAGATGTCGCGGTG 1075  
 DB 1574 TGACTTTGGTGGGCAATCATGCTGAGCTGGGCGCTGGGCTGTGGACAGAGATGTCGCGGTG 1633  
 QY 1076 TGTACACAAAGTTACCAACTACCTAGACTGATGATTCGTGACAACTGCGAGCG 1128  
 DB 1634 TGTACACAAAGTTACCAACTACCTAGACTGATGATTCGTGACAACTGCGAGCG 1686

RESULT 6

AAQ01358 standard; DNA; 1780 BP.

AC AAQ01358;

DT 25-MAR-2003 (updated)

DT 08-SEP-1990 (first entry)

DE Sequence encoding wild type tissue plasminogen activator (t-PA).

DX Wild type tissue plasminogen activator (t-PA); infarction treatment;

KW thrombosis treatment; embolism treatment.

XX Homo sapiens.

FT Key Location/Qualifiers

FT misc\_feature 1..123

FT /tag= a

FT /note= "synthetic 5' adaptor"

FT misc\_feature 1738..1780

FT /tag= b

FT /note= "synthetic 3' adaptor"

XX EP351246-A.

XX 17-JAN-1990.

XX 14-JUL-1989; 89EP-0307194.

XX 15-JUL-1988; 88DK-0003952.

XX (NOVO ) NOVO-NORDISK AS.

XX Petersen LC, Boel E;

XX

DR WPI: 1990-016567/03.

XX P-PSDB; AAR04699.

XX New tissue plasminogen activator (t-PA) analogue -  
 PT with higher fibrin selectivity than native t-PA, useful for  
 PT treating infarction, thrombosis and embolism

PS Disclosure; Fig 5A-D; 24pp; English.

XX It is modified in the patent by replacing one or more codons specifying  
 CC a positively charged amino acid. The resulting analogues are inserted  
 CC into a replicable expression vector which is used to transform or  
 CC transfect a host cell which is grown to express a t-PA analogue. The  
 CC t-PA analogue is useful for the treatment of diseases or disorders  
 CC associated with the formation of thrombi in blood vessels e.g.  
 CC infarctions, thrombosis and embolism. The analogue, in the 1-chain form,  
 CC exhibits the properties of a proenzyme; however on plasmin-catalysed  
 CC cleavage of the one-chain form, the activity of the 2-chain form is fully  
 CC retained. Compared to native t-PA, it has a higher fibrin selectivity,  
 CC this results in a fibrinolytic agent with a higher fibrin selectivity,  
 CC as the fibrinogenolytic activity induced by the analogue is reduced  
 CC relative to that induced by native 1-chain t-PA.  
 CC (Updated on 25-MAR-2003 to correct PA field.)

XX SQ Sequence 1780 BP; 405 A; 513 C; 523 G; 339 T; 0 other;

Query Match 94.4%; Score 1065; DB 11; Length 1780;

Best Local Similarity 99.5%; Pred. No. 4e-225;

Matches 1068; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 56 CCACAGGCGGCTCTGAGGGAACAGATGACTGCTACTTTGGGAATGGGTGACGCTTACCGTG 115  
 DB 630 CCCCTGCTCTCTGAGGGAACAGATGACTGCTACTTTGGGAATGGGTGACGCTTACCGTG 689  
 QY 116 GCACGCAACGCTCACCGAGTCGGGTGCTCTCTGCTCCCTGGAATTCATGATCTCTGA 175  
 DB 690 GCACGCAACGCTCACCGAGTCGGGTGCTCTCTGCTCCCTGGAATTCATGATCTCTGA 749  
 QY 176 TAGCAGAGTTTACACAGCAGACAGAACCCAGTGCAGGAGGCTGGGCTGGGCAACATA 235  
 DB 750 TAGCAGAGTTTACACAGCAGACAGAACCCAGTGCAGGAGGCTGGGCTGGGCAACATA 809  
 QY 236 ATTACTGCGGAATCTCTGATGGGATGCCAAGGCTCTGTCGACGCTGCAAGAACCGCA 295  
 DB 810 ATTACTGCGGAATCTCTGATGGGATGCCAAGGCTCTGTCGACGCTGCAAGAACCGCA 869  
 QY 296 GGCTGACGTGGGAGTACTGTGATGTGCTCTCTCTCCACTGCGGCTGAGACAGTACA 355  
 DB 870 GGCTGACGTGGGAGTACTGTGATGTGCTCTCTCTCCACTGCGGCTGAGACAGTACA 929  
 QY 356 GCACGCTCAGTTTCGATCAAAAGAGGCTCTTTCGCGGACATGCTCCACCCCTGGC 415  
 DB 930 GCACGCTCAGTTTCGATCAAAAGAGGCTCTTTCGCGGACATGCTCCACCCCTGGC 989  
 QY 416 AGCTGCTCATCTTTGCCAAGCAGAGAGGTCGCCCGGAGAGCGGTTCTGTGGGGGCA 475  
 DB 990 AGCTGCTCATCTTTGCCAAGCAGAGAGGTCGCCCGGAGAGCGGTTCTGTGGGGGCA 1049  
 QY 476 TACTCATCAGCTCCTGTGATTTCTCTCTGCGGCCACTGCTCCAGGAGAGGTTTCGCG 535  
 DB 1050 TACTCATCAGCTCCTGTGATTTCTCTCTGCGGCCACTGCTCCAGGAGAGGTTTCGCG 1109  
 QY 536 CCACACACCTGACGCTGATCTTGGGAGAACATACCGGGTGGTCCCTGGGAGAGAGAGC 595  
 DB 1110 CCACACACCTGACGCTGATCTTGGGAGAACATACCGGGTGGTCCCTGGGAGAGAGAGC 1169  
 QY 596 AGAAATTTGAAGTCGAAATATACATTGTCATAGGAATTCGATGACACTTACGACA 655  
 DB 1170 AGAAATTTGAAGTCGAAATATACATTGTCATAGGAATTCGATGACACTTACGACA 1229  
 QY 656 ATGACATTTGCGCTGCTGACACTGAAATTCGATTCGCTCCCGCTGTGCCAGAGAGAGCG 715

Db 1230 ATGACATTGGCGTGTGACGTGAATCGGATTCGTCCGCTGTGCCAGAGACGCG 1289  
Qy 716 TGGTCCGACATGTGTGCTTCCCCGGCGGACCTGACGTGCGGACTGGACGAGTGTG 775  
Db 1290 TGGTCCGACATGTGTGCTTCCCCGGCGGACCTGACGTGCGGACTGGACGAGTGTG 1349  
Qy 776 AGCTCTCCGCTTACGGCAAGCATGAGGCTTGTCTCTTCTATTCGGAGCGGCTGAAGG 835  
Db 1350 AGCTCTCCGCTTACGGCAAGCATGAGGCTTGTCTCTTCTATTCGGAGCGGCTGAAGG 1409  
Qy 836 AGGCTCATGTGACAGTGTACCATCCAGCCGCTGCACATCAACAATTTACTTTAACAGAA 895  
Db 1410 AGGCTCATGTGACAGTGTACCATCCAGCCGCTGCACATCAACAATTTACTTTAACAGAA 1469  
Qy 896 CAGTCACCGACAACATGTGTGTGTGTGAGACACTCGGAGCGGGGGCCCCAGGCAAACT 955  
Db 1470 CAGTCACCGACAACATGTGTGTGTGTGAGACACTCGGAGCGGGGGCCCCAGGCAAACT 1529  
Qy 956 TGCAGAGCGCTGCCAGGGCGATTCCGGAGGCGCCCTGGTGTCTGAACGATGCCCGCA 1015  
Db 1530 TGCAGAGCGCTGCCAGGGCGATTCCGGAGGCGCCCTGGTGTCTGAACGATGCCCGCA 1589  
Qy 1016 TGACTTTGGTGGGCATCATAGCTGGGCGCTGTGGACAGAGGATGTCCCGGGTG 1075  
Db 1590 TGACTTTGGTGGGCATCATAGCTGGGCGCTGTGGACAGAGGATGTCCCGGGTG 1649  
Qy 1076 TGTACAAAAGTTTACCAACTACCTAGACTGTGATTCGTGACAAATGCGACCG 1128  
Db 1650 TGTACAAAAGTTTACCAACTACCTAGACTGTGATTCGTGACAAATGCGACCG 1702

## RESULT 7

ID AAN91119 standard; DNA; 2100 BP.

AC AAN91119;

XX 25-MAR-2003 (updated)

DT 03-OCT-2002 (updated)

DT 18-JUN-1990 (first entry)

XX

XX Sequence of native tPA in plasmid pST112.

XX Tissue plasminogen activator; tPA; thrombolytic agent;

KW plasminogen; vascular diseases.

XX Synthetic.

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

PS Example 29; Fig 21; 68pp; English.

XX cDNA sequence of native tPA gene is excised from plasmid pST112, and

CC digested with BamHI and SalI to form plasmid pST118.

CC (Updated on 03-OCT-2002 to add missing OS field.)

CC (Updated on 25-MAR-2003 to correct PA field.)

XX

SQ Sequence 2100 BP; 498 A; 595 C; 597 G; 410 T; 0 other;

Query Match 94.4%; Score 1065; DB 10; Length 2100;

Best Local Similarity 99.5%; Pred. No. 4.1e-225;

Matches 1068; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 56 CCCAGCGCGCTCTGAGGGAACAGTGTACTTTGGGAATGGGTACGCTTACCGTG 115

Db 638 CCCCTGCTGCTCTGAGGGAACAGTGTACTTTGGGAATGGGTACGCTTACCGTG 697

Qy 116 GCAGCAGCAGCTCACCAGGTGGGTGCTCTCTCCCTCCGCTGGAATTCATGATCCTGA 175

Db 698 GCAGCAGCAGCTCACCAGGTGGGTGCTCTCTCCCTCCGCTGGAATTCATGATCCTGA 757

Qy 176 TAGCAAGGTTTACACAGCAGAACCCAGTGCAGGCACTGGGCTGGGCAAAATA 235

Db 758 TAGCAAGGTTTACACAGCAGAACCCAGTGCAGGCACTGGGCTGGGCAAAATA 817

Qy 236 ATTACTCCGGAATCCTGATGGGATGCCAAGCCCTGGTGCACGTGCTGAAGAACCGCA 295

Db 818 ATTACTCCGGAATCCTGATGGGATGCCAAGCCCTGGTGCACGTGCTGAAGAACCGCA 877

Qy 296 GGCTGAGTGGGAGTACTGTGATGTGCTCTCTCTCCCTGACCTGGGCTGAGACAGTACA 355

Db 878 GGCTGAGTGGGAGTACTGTGATGTGCTCTCTCTCCCTGACCTGGGCTGAGACAGTACA 937

Qy 356 GCCAGCTCAGTTTCGATCAAGAGGAGGTCTTTCGCGACATCGCTCCACCCCTGGC 415

Db 938 GCCAGCTCAGTTTCGATCAAGAGGAGGTCTTTCGCGACATCGCTCCACCCCTGGC 997

Qy 416 AGGCTGCATCTTTGCCAAGCAGAGGAGTGCCTCCGCGAGAGCGGTTCCTGTCGGGGCA 475

Db 998 AGGCTGCATCTTTGCCAAGCAGAGGAGTGCCTCCGCGAGAGCGGTTCCTGTCGGGGCA 1057

Qy 476 TACTCATCAGCTCTGCTGATTTCTCTGCGCGCCACTGCTTCCAGGAGAGGTTTCGCG 535

Db 1058 TACTCATCAGCTCTGCTGATTTCTCTGCGCGCCACTGCTTCCAGGAGAGGTTTCGCG 1117

Qy 536 CCCACACCTGACGCTGATCTTGGGAGAACATACCGGTGCTCCCTGGCGAGAGGAGC 595

Db 1118 CCCACACCTGACGCTGATCTTGGGAGAACATACCGGTGCTCCCTGGCGAGAGGAGC 1177

Qy 596 AGAAATTTGAAGTCGAAAAATACATTTGTCCATAAGGAATTCGATGATGACACTTACGACA 655

Db 1178 AGAAATTTGAAGTCGAAAAATACATTTGTCCATAAGGAATTCGATGATGACACTTACGACA 1237

Qy 656 ATGACATTCGCTCTGCTGAGTGAATTCGATTCGCTCCGCTGTGCCAGAGAGCAGCG 715

Db 1238 ATGACATTCGCTCTGCTGAGTGAATTCGATTCGCTCCGCTGTGCCAGAGAGCAGCG 1297

Qy 716 TGGTCCGCACTGTGTGCTTCCCGCGGACCTGACGCTGCCGACTGGACGAGTGTG 775

Db 1298 TGGTCCGCACTGTGTGCTTCCCGCGGACCTGACGCTGCCGACTGGACGAGTGTG 1357

Qy 776 AGCTCTCCGCTACGCAAGCATGAGGCTTGTCTCTCTTTCTATTTCGAGCGGCTGAAGG 835

Db 1358 AGCTCTCCGCTACGCAAGCATGAGGCTTGTCTCTCTTTCTATTTCGAGCGGCTGAAGG 1417

Qy 836 AGGCTCATGTGACAGTGTATCCCATCCAGCGCTGACATCAACAATTTACTTTAACAGAA 895

Db 1418 AGGCTCATGTGACAGTGTATCCCATCCAGCGCTGACATCAACAATTTACTTTAACAGAA 1477

Qy 896 CAGTCACCGACAACATCTGTGTGCTGAGACACTCGGAGCGGGGGCCCCAGGCAAACT 955

Db 1478 CAGTCACCGACAACATCTGTGTGCTGAGACACTCGGAGCGGGGGCCCCAGGCAAACT 1537

QY 956 TGCACGAGCCTGCCAGGCGGATTCGGAGGCCCCCTGTGTGTCTGAACGATGCCGCA 1015  
 DB 1538 TGCACGAGCCTGCCAGGCGGATTCGGAGGCCCCCTGTGTGTCTGAACGATGCCGCA 1597  
 QY 1016 TGACTTTGGTGGGCAATCATCAGCTGGGCGCTGGGCTGTGGACAGAGGATGTCCGGGTG 1075  
 DB 1598 TGACTTTGGTGGGCAATCATCAGCTGGGCGCTGGGCTGTGGACAGAGGATGTCCGGGTG 1657  
 QY 1076 TGTACACAAAGGTTACCAACTACCTAGACTGGATTTCGTGACAAACATGCGACG 1128  
 DB 1658 TGTACACAAAGGTTACCAACTACCTAGACTGGATTTCGTGACAAACATGCGACG 1710

## RESULT 8

AAQ05532  
 ID AAQ05532 standard; DNA; 2100 BP.

XX

AC AAQ05532;

XX

DT 25-MAR-2003 (updated)

DT 11-DEC-1990 (first entry)

XX

XX Plasmid pStr12 encoding novel N-terminal for tissue plasminogen

DE activator (tPA).

XX

KW Fibrin; clotting; thrombolytic; vascular disease; stroke;

KW myocardial infarction; heart attack; pulmonary embolism; ds;

XX

XX Homo sapiens.

XX

XX Key

FT CDS

FT

FT

FT

FT

FT

XX

PN EP379890-A.

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

QY 116 GCACGACAGCCTCACCGAGTCGGGTGCTCTCTGCTCCCTCCCTGGAATTCATGATCCTGA 175  
 DB 698 GCACGACAGCCTCACCGAGTCGGGTGCTCTCTGCTCCCTCCCTGGAATTCATGATCCTGA 757  
 QY 176 TAGGCAAGGTTTACACAGCACAGAAACCCAGTGCCTCAGGCACTGGGCTGGGCAAAACATA 235  
 DB 758 TAGGCAAGGTTTACACAGCACAGAAACCCAGTGCCTCAGGCACTGGGCTGGGCAAAACATA 817  
 QY 236 ATTACTCCGGAATCTCATGCGGATGCCAAGCCCTGGTGCACGTCGTGAAGACCCGCA 295  
 DB 818 ATTACTCCGGAATCTCATGCGGATGCCAAGCCCTGGTGCACGTCGTGAAGACCCGCA 877  
 QY 296 GGCTGAGCTGGGAGTACTGTGATGCTCTCTCTGCTCCCTCCCTGCGGCTGAGACAGTACA 355  
 DB 878 GGCTGAGCTGGGAGTACTGTGATGCTCTCTCTGCTCCCTCCCTGCGGCTGAGACAGTACA 937  
 QY 356 GCCAGCCTCAGTTTGGCATCAAGAGGAGGCTCTTTGCGGCAATCGCCTCCCAACCCCTGGC 415  
 DB 938 GCCAGCCTCAGTTTGGCATCAAGAGGAGGCTCTTTGCGGCAATCGCCTCCCAACCCCTGGC 997  
 QY 416 AGGCTGCCATCTTTGCCAAGCACAGGAGGTGCGCGGAGAGCGGTTCTGTGCGGCGCA 475  
 DB 998 AGGCTGCCATCTTTGCCAAGCACAGGAGGTGCGCGGAGAGCGGTTCTGTGCGGCGCA 1057  
 QY 476 TACTCATCAGCTCCTGTGATTCCTCTGCGGCCACTGCTTCAGGAGAGGTTTCCGC 535  
 DB 1058 TACTCATCAGCTCCTGTGATTCCTCTGCGGCCACTGCTTCAGGAGAGGTTTCCGC 1117  
 QY 536 CCCACCACTGACGCTGATCTTTGGGCAAAACATACCGGGTGTCTCTGCGGAGAGGAGC 595  
 DB 1118 CCCACCACTGACGCTGATCTTTGGGCAAAACATACCGGGTGTCTCTGCGGAGAGGAGC 1177  
 QY 596 AGAATTTGAGTCGAAATACATTTCCATAGGAATTCGATGATGACACTTACGACA 655  
 DB 1178 AGAATTTGAGTCGAAATACATTTCCATAGGAATTCGATGATGACACTTACGACA 1237  
 QY 656 ATGACATTTGGCTGCTGACGCTGAAATCGGATTCCTCCGCTGTCGCCAGAGAGCAGCG 715  
 DB 1238 ATGACATTTGGCTGCTGACGCTGAAATCGGATTCCTCCGCTGTCGCCAGAGAGCAGCG 1297  
 QY 716 TGGTCCGACCTGTGTGCTTCCCGCGGACCTGACGCTGCGGAGTGGACGAGTGTG 775  
 DB 1298 TGGTCCGACCTGTGTGCTTCCCGCGGACCTGACGCTGCGGAGTGGACGAGTGTG 1357  
 QY 776 AGCTCTCGGCTAGCGCAAGCATGAGGCTGTCTCTCTCTTCTTATTCGAGGCGGCTGAGG 835  
 DB 1358 AGCTCTCGGCTAGCGCAAGCATGAGGCTGTCTCTCTCTTCTTATTCGAGGCGGCTGAGG 1417  
 QY 836 AGGCTCATGTGACACTGTACCCATCCAGCGCTGCACATCACAACTTTACTTAACAGAA 895  
 DB 1418 AGGCTCATGTGACACTGTACCCATCCAGCGCTGCACATCACAACTTTACTTAACAGAA 1477  
 QY 896 CAGTACCGCAACACATCTGTGTGCTGAGACACTCGGAGCGGCGGCCCCAGGCAAACT 955  
 DB 1478 CAGTACCGCAACACATCTGTGTGCTGAGACACTCGGAGCGGCGGCCCCAGGCAAACT 1537  
 QY 956 TGACAGAGCGCTGCGAGGCGATTCGGAGGCGCCCTGCTGTCTGACAGTGGCGGCA 1015  
 DB 1538 TGACAGAGCGCTGCGAGGCGATTCGGAGGCGCCCTGCTGTGTGTCGAACGATGGCGCA 1597  
 QY 1016 TGACTTTGGTGGGCAATCATCAGCTGGGCGCTGGGCTGTGGACAGAGGATGTCCCGGCTG 1075  
 DB 1598 TGACTTTGGTGGGCAATCATCAGCTGGGCGCTGGGCTGTGGACAGAGGATGTCCCGGCTG 1657  
 QY 1076 TGATACAAAGGTTTACCACTAGACTGGATTCTGTGACCAACATGCGACCG 1128  
 DB 1658 TGTACACAAAGGTTTACCACTAGACTGGATTCTGTGACCAACATGCGACCG 1710

## RESULT 9

AAQ87370

ID AAQ87370 standard; DNA; 2162 BP.

XX

Query Match 94.4%; Score 1065; DB 11; Length 2100;  
 Best Local Similarity 99.5%; Pred. No. 4.1e-225;  
 Matches 1068; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 56 CCCAGCGGCTCTGAGGGAACAGTACTGCTACTTTGGGAATGGGTACGCTACCGTG 115

DB 638 CCCCTGCTCTGAGGGAACAGTACTGCTACTTTGGGAATGGGTACGCTACCGTG 697

PT New tissue plasminogen activator - having N-terminal peptide of  
 PT plasminogen linked to tissue plasminogen activator for increased  
 PT stability in vivo.

PS Disclosure; Page ?; ?pp; English.

XX tPA with N-terminal peptide of plasminogen linked is more stable in.  
 CC vivo than the native form. It is useful as a thrombolytic agent in  
 CC the treatment of vascular diseases eg myocardial infarction,  
 CC pulmonary embolism etc.  
 CC (Updated on 25-MAR-2003 to correct PA field.)

SQ Sequence 2100 BP; 498 A; 595 C; 597 G; 410 T; 0 other;



AC AAQ87370;  
 XX 25-MAR-2003 (updated)  
 DT 19-SEP-1995 (first entry)  
 XX  
 DE Human tPA clone.  
 XX  
 KW Intron; recombination; combinatorial gene; trans-splicing;  
 KW gene therapy; polymerase chain reaction; PCR; primer; amplification;  
 KW tissue plasminogen activator; tPA; plasmid tPA-KS+; thrombolytic;  
 KW ds.  
 XX  
 OS Synthetic.  
 XX  
 FH Key Location/Qualifiers  
 FT sig\_peptide 82..334  
 FT /tag= a  
 FT /note= "signal sequence and finger-like domain"  
 FT CDS 335..447  
 FT /tag= b  
 FT /product= EGF-like domain  
 FT CDS 448..714  
 FT /tag= c  
 FT /product= Kringle-1 domain  
 FT CDS 715..972  
 FT /tag= d  
 FT /product= Kringle-2 domain  
 FT CDS 973..2162  
 FT /tag= e  
 FT /product= catalytic domain  
 XX  
 PN W09507351-Al.  
 XX  
 PD 16-MAR-1995.  
 XX  
 PP 12-SEP-1994; 94WO-US10146.  
 XX  
 PR 10-SEP-1993; 93US-0119512.  
 XX  
 PA (HARD ) HARVARD COLLEGE.  
 XX  
 PI Jarrell KA;  
 XX  
 DR WPI; 1995-123425/16.  
 XX  
 PT New intron-mediated recombinant techniques - used for the  
 PT generation and selection of novel genes and gene prods. for use  
 PT in therapy  
 XX  
 PS Example 4; Page 59-60; 87pp; English.  
 XX  
 CC A cDNA clone of human tissue plasminogen activator (tPA) was  
 CC amplified by PCR using the primers given in AAQ87368-69. The  
 CC amplified tPA DNA (AAQ87370) was ligated into vector KS+ to  
 CC obtain plasmid tPA-KS+. The construct was used in combinatorial  
 CC methods involving RNA splicing-mediated shuffling of tPA domains  
 CC in plasmid PINVI (AAQ87347) to generate novel tPAs having  
 CC improved thrombolytic properties.  
 CC (Updated on 25-MAR-2003 to correct PN field.)  
 XX  
 SQ Sequence 2162 BP; 520 A; 606 C; 619 G; 417 T; 0 other;  
 Query Match 94.4%; Score 1065; DB 16; Length 2162;  
 Best Local Similarity 99.5%; Pred. No. 4.1e-225;  
 Matches 1068; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
 QY 56 CCCAGCGGCTCTGAGGAAACAGTGTCTACTTTGGGAATGGGTACGCTACCGTG 115  
 DB 695 CCCCTGCTGTCTGAGGAAACAGTGTCTACTTTGGGAATGGGTACGCTACCGTG 754  
 QY 116 GCACGACAGCTTACCGAGTCGGGTGCTCTGCTCCGCTGGAAATCCATGATCTTGA 175  
 DB 755 GCACGACAGCTTACCGAGTCGGGTGCTCTGCTCCGCTGGAAATCCATGATCTTGA 814

QY 176 TAGCAAGTTTACACAGCACAGAAACCCAGTCCCGAGGCACTGGCTGGGCAAAACATA 235  
 DB 815 TAGCAAGTTTACACAGCACAGAAACCCAGTCCCGAGGCACTGGCTGGGCAAAACATA 874  
 QY 236 ATTACTCCGGAATCCTGATGGGGATCCCAAGCCCTGGTGCACAGTGTGAGACCGCA 295  
 DB 875 ATTACTCCGGAATCCTGATGGGGATCCCAAGCCCTGGTGCACAGTGTGAGAAACCGCA 934  
 QY 296 GGCTGAGCTGGGAGTACTGTGATGTGCCCTCTCTCCACCTGGGGCTGAGACAGTACA 355  
 DB 935 GGCTGAGCTGGGAGTACTGTGATGTGCCCTCTCTCCACCTGGGGCTGAGACAGTACA 994  
 QY 356 GCCAGCTCAGTTTCCGATCAAGAGGGCTCTTTCGCCGACATCGCTCCCAACCCCTGGC 415  
 DB 995 GCCAGCTCAGTTTCCGATCAAGAGGGCTCTTTCGCCGACATCGCTCCCAACCCCTGGC 1054  
 QY 416 AGGCTGCCATCTTTGCCAAGCAGAGGGTCCCGGGAGAGGGTTCCTGTGGGGGCA 475  
 DB 1055 AGGCTGCCATCTTTGCCAAGCAGAGGGTCCCGGGAGAGGGTTCCTGTGGGGGCA 1114  
 QY 476 TACTCATCAGCTCTGTGATTTCTCTGCCGCCACTGTCTTCAGAGAGGTTTTCGCG 535  
 DB 1115 TACTCATCAGCTCTGTGATTTCTCTGCCGCCACTGTCTTCAGAGAGGTTTTCGCG 1174  
 QY 536 CCCACCACTGACCGGTGATTTGGGCAAAACATACCGGGTGGTCCCTGGCGAGAGGAGC 595  
 DB 1175 CCCACCACTGACCGGTGATTTGGGCAAAACATACCGGGTGGTCCCTGGCGAGGAGC 1234  
 QY 596 AGAATTTGAGTCGAAATACATTTGCCATAGGAATTCGATGATGACACTTACGCA 655  
 DB 1235 AGAATTTGAGTCGAAATACATTTGCCATAGGAATTCGATGATGACACTTACGCA 1294  
 QY 656 ATGACATTTGCTGCTGAGCTGAAATCGAATTCGTCCCGCTGTGCCCGAGAGCAGCG 715  
 DB 1295 ATGACATTTGCTGCTGAGCTGAAATCGAATTCGTCCCGCTGTGCCCGAGAGCAGCG 1354  
 QY 716 TGGTCCGCACTGTGTGCTTCCCGGGGACCTTCAGCTCCCGAGTGGACGAGTGTG 775  
 DB 1355 TGGTCCGCACTGTGTGCTTCCCGGGGACCTTCAGCTCCCGAGTGGACGAGTGTG 1414  
 QY 776 AGCTCTCGGCTACGGCAAGCATGAGGCTTGTCTCTCTTCTATTCGAGCGGCTGAAG 835  
 DB 1415 AGCTCTCGGCTACGGCAAGCATGAGGCTTGTCTCTCTTCTATTCGAGCGGCTGAAG 1474  
 QY 836 AGGCTCATGTGAGCTGTACCCATCCAGCTGCACATCACAACATTTACTTACAGAA 895  
 DB 1475 AGGCTCATGTGAGCTGTACCCATCCAGCTGCACATCACAACATTTACTTACAGAA 1534  
 QY 896 CAGTCACCGCAACATGCTGTGTCTGGAGACACTCGGAGCGGGGCCCCAGGCAAACT 955  
 DB 1535 CAGTCACCGCAACATGCTGTGTCTGGAGACACTCGGAGCGGGGCCCCAGGCAAACT 1594  
 QY 956 TGCACGAGCTTGCAGGGCGATTCGGAGGCGCCCTGGTGTGTGTAACATGGCCGCA 1015  
 DB 1595 TGCACGAGCTTGCAGGGCGATTCGGAGGCGCCCTGGTGTGTGTAACATGGCCGCA 1654  
 QY 1016 TGACTTTGGTGGGATCATCAGCTGGGCTGTGGACAGAGGATGTCCCGGGTG 1075  
 DB 1655 TGACTTTGGTGGGATCATCAGCTGGGCTGTGGACAGAGGATGTCCCGGGTG 1714  
 QY 1076 TGTACAAAGGTTTACCAACTACCTAGACTGGATTCGTGACAAACATGCGACCG 1128  
 DB 1715 TGTACAAAGGTTTACCAACTACCTAGACTGGATTCGTGACAAACATGCGACCG 1767

RESULT 10  
 AAV37294  
 ID AAV37294 standard; DNA; 2162 BP.  
 XX  
 AC AAV37294;  
 XX  
 DT 10-SEP-1998 (first entry)

XX Human tissue plasminogen activator gene sequence.

DE Plasmid pINVI; reverse-splicing intron; group II intron;

XX exon binding site; domain V motif; branch site acceptor;

KW nucleophilic group; transesterification; phosphodiester bond;

KW autocatalytic Y-branched intron; reverse splicing reaction; ds.

XX Homo sapiens.

OS

XX

PH Key Location/Qualifiers

FT misc\_feature 82..334

FT /tag= a

FT /note= "signal sequence and finger-like domain"

FT misc\_feature 335..447

FT /tag= b

FT /note= "EGF-like domain"

FT misc\_feature 448..714

FT /tag= c

FT /note= "Kringle-1 domain"

FT misc\_feature 715..972

FT /tag= d

FT /note= "Kringle-2 domain"

FT misc\_feature 973..2162

FT /tag= e

FT /note= "catalytic domain"

XX

PN US5780272-A.

XX

XX 14-JUL-1998.

XX

XX 07-JUN-1995; 95US-0488015.

XX

PR 10-SEP-1993; 93US-0119512.

XX

XX (HARD ) HARVARD COLLEGE.

XX

XX Jarrell KA;

XX

DR WPI; 1998-413060/35.

XX

XX Reverse splicing construct containing fragments of autocatalytic

PT introns - able to cleave and ligate discontinuous nucleic acid for

PT generating new genes and e.g. ribozymes, libraries of enzymes and

PT antibodies

XX

PS Example 4; Columns 53-56; 56pp; English.

XX

XX The present sequence represents the human tissue plasminogen activator

CC gene. It was used to construct plasmid TPA-KS+, which is used in the

CC course of the invention. The specification describes a purified

CC reverse-splicing intron which comprises a segment comprising a

CC 5'-part of a group II intron, including an exon binding site not

CC naturally present in the intron and a second segment comprising a

CC 3'-part of a group II intron, including a domain V motif, a branch

CC site acceptor, and a nucleophilic group for transesterifying a

CC phosphodiester bond of an RNA. Together the two segments form an

CC autocatalytic Y-branched intron which catalyses integration of at least

CC the first segment into substrate RNA by a reverse splicing reaction

CC The reverse-splicing introns are used, by specific cleavage and ligation

CC of discontinuous nucleic acid, to generate new genes and gene products,

CC e.g. ribozymes (for use in gene therapy or as reagents in DNA

CC manipulation, e.g. replacements for restriction enzymes) or

CC immunologically active or signal-transducing proteins such as antibody

CC and enzyme libraries.

XX

SQ Sequence 2162 BP; 520 A; 606 C; 619 G; 417 T; 0 other;

XX

Query Match 94.4%; Score 1065; DB 19; Length 2162;

Best Local Similarity 99.5%; Fred. No. 4.1e-225;

Matches 1068; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

XX

56 CCCAGCGCGCTCTGAGGAAACAGTACTGTCTACTTTGGGAATGGGTACGCTACCGTG 115

695 CCCCTGCTCTCTGAGGGAAACAGTACTGTCTACTTTGGGAATGGGTACGCTACCGTG 754

116 GCAGCAGAGCTCACCAGTGGGTGCTCTGCTCCCTCCCGTGAATTCATGATCCTGA 175

755 GCAGCAGAGCTCACCAGTGGGTGCTCTGCTCCCTCCCGTGAATTCATGATCCTGA 814

176 TAGCAAGGTTTACACAGCAGAACCCAGTGGGTGCTCTGCTCCCGTGAATTCATGAT 235

815 TAGCAAGGTTTACACAGCAGAACCCAGTGGGTGCTCTGCTCCCGTGAATTCATGAT 874

236 ATTACTCCCGGAATCCTGATGGGATGCCAAGCCCTGGTGCACGTGCTGAAGAACCGCA 295

875 ATTACTCCCGGAATCCTGATGGGATGCCAAGCCCTGGTGCACGTGCTGAAGAACCGCA 934

296 GGCTGAGCTGGAGTACTGTGATGTGCCCTCTGCTCCACTGCGGCTGAGACAGTACA 355

935 GGCTGAGCTGGAGTACTGTGATGTGCCCTCTGCTCCACTGCGGCTGAGACAGTACA 994

356 GCCAGCCTCAGTTCGCATCAAGAGGAGGCTCTTCGCCGACATCGCTCCACCCCTGGC 415

995 GCCAGCCTCAGTTCGCATCAAGAGGAGGCTCTTCGCCGACATCGCTCCACCCCTGGC 1054

416 AGGCTGCCATCTTTGCCAAGCAGAGAGGTGCGCCGAGAGCGGTTCCTGTGCGGGGCA 475

1055 AGGCTGCCATCTTTGCCAAGCAGAGAGGTGCGCCGAGAGCGGTTCCTGTGCGGGGCA 1114

476 TACTCATCAGCTCCTGCTGGATTCCTCTGCGGCCCTCTGCTCCAGAGAGGTTTCGCG 535

1115 TACTCATCAGCTCCTGCTGGATTCCTCTGCGGCCCTCTGCTCCAGAGAGGTTTCGCG 1174

536 CCCACCACTGAGCGTGTCTTGGGAGAGACATACCGGTGGTCCCTGGCGAGGAGGAGC 595

1175 CCCACCACTGAGCGTGTCTTGGGAGAGACATACCGGTGGTCCCTGGCGAGGAGGAGC 1234

596 AGAAATTTGAAGTCGAAAAATACATTTCTCCATAGGAATTCGATGATGACACTTACGACA 655

1235 AGAAATTTGAAGTCGAAAAATACATTTCTCCATAGGAATTCGATGATGACACTTACGACA 1294

656 ATGACATTTGGCTGCTGCGAGTGAATTCGATTCGCTCCCGTGTGCCAGAGAGCAGCG 715

1295 ATGACATTTGGCTGCTGCGAGTGAATTCGATTCGCTCCCGTGTGCCAGAGAGCAGCG 1354

716 TGGTCCGACTGTGTGCTTCCCGCGGAGACCTGCGAGCTGCGGAGTGGACGAGTGTG 775

1355 TGGTCCGACTGTGTGCTTCCCGCGGAGACCTGCGAGCTGCGGAGTGGACGAGTGTG 1414

776 AGCTCTCCGGCTACGCAAGCATGAGGCTTGTCTCTCTTTCTATTTCGAGCGGCTGAAGG 835

1415 AGCTCTCCGGCTACGCAAGCATGAGGCTTGTCTCTCTTTCTATTTCGAGCGGCTGAAGG 1474

836 AGGCTCATGTGAGTGTATCCATTCAGCGCTGCTGCATACATCAACATTTACTTAACAGAA 895

1475 AGGCTCATGTGAGTGTATCCATTCAGCGCTGCTGCATACATCAACATTTACTTAACAGAA 1534

896 CAGTCACCGCAACATCTGTGTCTGAGACACTCGAGCGGCGGCCCCAGAGCAAACT 955

1535 CAGTCACCGCAACATCTGTGTCTGAGACACTCGAGCGGCGGCCCCAGAGCAAACT 1594

956 TGCAACGAGCGCTGCGAGCGGATTCGGGAGCGGCCCTTGGTGTGTCTGAACGATGCGCGCA 1015

1595 TGCAACGAGCGCTGCGAGCGGATTCGGGAGCGGCCCTTGGTGTGTCTGAACGATGCGCGCA 1654

1016 TGACTTTGGTGGGATCATCATGCTGGGCTTGGGCTGTGGACAGAGATGTCGCGGGTG 1075

1655 TGACTTTGGTGGGATCATCATGCTGGGCTTGGGCTGTGGACAGAGATGTCGCGGGTG 1714

1076 TGATACAAAGGTTTACCAACTTACCTAGACTGGATTCTGTGACAACTGCGACG 1128

1715 TGATACAAAGGTTTACCAACTTACCTAGACTGGATTCTGTGACAACTGCGACG 1767

RESULT 11

AAN60659  
ID AAN60659 standard; DNA; 2547 BP.  
XX  
AC AAN60659;  
XX  
DT 25-MAR-2003 (updated)  
DT 22-JUL-1991 (first entry)  
XX  
DE Sequence encoding human pre-tissue plasminogen activator (pre-t-PA).  
XX  
XX Plasminogen conversion; vascular disease therapy; ss.  
XX  
XX Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT CDS 87..191  
FT FT /\*tag= a  
FT mat\_peptide 192..1775  
FT FT /\*tag= b  
XX  
XX GB2173804-A.  
XX  
XX PD 22-OCT-1986.  
XX  
XX PF 21-APR-1986; 86GB-0609683.  
XX  
XX PR 22-APR-1985; 85US-0725468.  
PR 01-APR-1986; 86US-0846697.  
PR 22-APR-1985; 85US-0725468.  
PR 01-APR-1986; 86US-0846697.  
XX  
XX (GETH) GENENTECH INC.  
XX  
XX Heyneker HL, Vehar GA;  
XX  
XX WPI: 1986-280715/43.  
XX P-PSDB; AAP60790.  
XX  
XX New mutant forms of human tissue plasminogen activator- having  
XX higher specific activity and resistance to conversion to two chain  
XX form  
XX  
XX Example; Fig 2; 34pp; English.  
XX  
XX The patentors claim a novel recombinant human t-PA which is  
XX resistant to specific enzymatic cleavage because it is stabilised by  
XX site-directed mutagenesis at a 2-chain cleavage site. Partic. the  
XX natural Arg at position 275 is replaced by Gly or Glu, or Ile at  
XX position 276 is replaced (275 and 276 refer to the posn. of the AAs  
XX in the mature protein; i.e. AAs 310 and 311 of AAP60790). Also new are  
XX (1) DNA sequences; (2) expression vectors; and (3) microorganisms  
XX and cell cultures transformed with these vectors.  
XX (Updated on 25-MAR-2003 to correct PR field.)  
XX  
XX Sequence 2547 BP; 627 A; 688 C; 691 G; 541 T; 0 other;  
XX  
XX Query Match 94.4%; Score 1065; DB 7; Length 2547;  
XX Best Local Similarity 99.5%; Pred. No. 4.2e-225;  
XX Matches 1068; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
XX  
XX 56 CCCAGCGGCTCTGAGGAAACAGTGTCTCTCTGCTCCGCTGGGAAATTCATGATCCTGA 175  
XX  
XX 700 CCCCTGCTGTCTGAGGAAACAGTGTCTCTCTGCTCCGCTGGGAAATTCATGATCCTGA 759  
XX  
XX 116 GCACGCAGCCTCACCGAGTGGGTGCTCTCTGCTCCGCTGGGAAATTCATGATCCTGA 175  
XX  
XX 760 GCACGCAGCCTCACCGAGTGGGTGCTCTCTGCTCCGCTGGGAAATTCATGATCCTGA 819  
XX  
XX 176 TAGGCAAGGTTTACACAGCAGAAACCCAGTGGCCAGGCAGTGGGCTGGGCAACATA 235  
XX  
XX 820 TAGGCAAGGTTTACACAGCAGAAACCCAGTGGCCAGGCAGTGGGCTGGGCAACATA 879  
XX  
XX 236 ATTACTGCGGAAATCTGTATGGGGATGCCAAGCCCTGGTGGCCAGTGTCTGAAGACCGCA 295

Db 880 ATTACTGCGGAAATCTGTATGGGGATGCCAAGCCCTGGTGGCCAGTGTCTGAAGAACCGCA 939  
Qy 296 GGCTGACGTGGAGTACTGTGTCTCTGTCTCCACTTGGCGGCTTGAGACAGTACA 355  
Db 940 GGCTGACGTGGAGTACTGTGTCTCTGTCTCCACTTGGCGGCTTGAGACAGTACA 999  
Qy 356 GCCAGCTCAGTTTTCGCATCAAGAGGGCTCTTTGGCCGACATGCCCTCCACCCCTGGC 415  
Db 1000 GCCAGCTCAGTTTTCGCATCAAGAGGGCTCTTTGGCCGACATGCCCTCCACCCCTGGC 1059  
Qy 416 AGGCTGCCATCTTTGCCAAGCAGAGAGGTGCCCGGAGAGCGGTTCCTGTGCGGGGCA 475  
Db 1060 AGGCTGCCATCTTTGCCAAGCAGAGAGGTGCCCGGAGAGCGGTTCCTGTGCGGGGCA 1119  
Qy 476 TACTCATCAGCTCTCTGTGGATTCTCTGTGCGGCCACTGTCTTCAGAGAGGTTTTCGCG 535  
Db 1120 TACTCATCAGCTCTCTGTGGATTCTCTGTGCGGCCACTGTCTTCAGAGAGGTTTTCGCG 1179  
Qy 536 CCCACCACTGACGGTGTCTTTGGGCAAAATACATCCGGGTGGTCCCTGGCGAGGAGGAGC 595  
Db 1180 CCCACCACTGACGGTGTCTTTGGGCAAAATACATCCGGGTGGTCCCTGGCGAGGAGGAGC 1239  
Qy 596 AGAAATTTGAAGTGCAGAAATACATTTGCCATAGGAATTCGATGATGACACTTACGACA 655  
Db 1240 AGAAATTTGAAGTGCAGAAATACATTTGCCATAGGAATTCGATGATGACACTTACGACA 1299  
Qy 656 ATGACATTTGGCTCTGTCAGCTGAAATCGGAATTCGTCCTGCTGCGGAGAGGAGGAGC 715  
Db 1300 ATGACATTTGGCTCTGTCAGCTGAAATCGGAATTCGTCCTGCTGCGGAGAGGAGGAGC 1359  
Qy 716 TGGTCCGCACTGTGTGCTTCCCGCGGCAACATCCAGCTGCGGAGTGGAGCGAGTGTG 775  
Db 1360 TGGTCCGCACTGTGTGCTTCCCGCGGCAACATCCAGCTGCGGAGTGGAGCGAGTGTG 1419  
Qy 776 AGCTCTCGGCTACGGCAAGCATGAGGCTTGTCTCTCTTCTATTCGAGGCGGCTGAAGG 835  
Db 1420 AGCTCTCGGCTACGGCAAGCATGAGGCTTGTCTCTCTTCTATTCGAGGCGGCTGAAGG 1479  
Qy 836 AGGCTCATGTGAGACTGTACCCATCCAGCTGACATCAACATTTACTTAAACAGAA 895  
Db 1480 AGGCTCATGTGAGACTGTACCCATCCAGCTGACATCAACATTTACTTAAACAGAA 1539  
Qy 896 CAGTCAACGCAACATGTGTGTGTGAGACACTTCGGAGCGGGGCGGCGGCAAACT 955  
Db 1540 CAGTCAACGCAACATGTGTGTGTGAGACACTTCGGAGCGGGGCGGCGGCAAACT 1599  
Qy 956 TGCACGACGCTGCGCAGGCGAATTCGGAGGCGGCGGCTGTGTGTGTGAACGATGGCCGCA 1015  
Db 1600 TGCACGACGCTGCGCAGGCGAATTCGGAGGCGGCGGCTGTGTGTGTGAACGATGGCCGCA 1659  
Qy 1016 TGACTTTGGTGGGCATCATCAGCTGGGCGCTGGGCTGGGACAGAGATGTCCGGGTG 1075  
Db 1660 TGACTTTGGTGGGCATCATCAGCTGGGCGCTGGGCTGGGACAGAGATGTCCGGGTG 1719  
Qy 1076 TGATACAAAGGTTTACCAACTTACCTAGACTGGATTCGTCGACAACTGCGACCG 1128  
Db 1720 TGATACAAAGGTTTACCAACTTACCTAGACTGGATTCGTCGACAACTGCGACCG 1772  
XX  
XX RESULT 12  
XX AAN90542  
XX ID AAN90542 standard; cDNA; 2560 BP.  
XX  
XX AC AAN90542;  
XX  
XX 25-MAR-2003 (updated)  
XX 05-JUN-1990 (first entry)  
XX  
XX Plasmid pKG12 contg. DNA encoding human melanoma t-PA.  
XX Human tissue plasminogen-activator gene; pKG12; ss.  
XX





```

RESULT 14
AA081970 standard; DNA; 7533 BP.
AC AA081970;
DT 12-OCT-1990 (first entry)
DE Plasmid pEMpl-TPA.
KW Tissue plasminogen activator; myocardial infarction; thrombosis;
KW embolism; ss.
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX misc_RNA 7..1748
XX FT /*tag= a
XX FT /label=tpa_cdna
XX FT 20..1708
XX FT /*tag= b
XX FT /product=t-PA
XX FT 1749..1955
XX FT /*tag= c
XX FT /label=k3'UT
XX FT /note="from Ig kappa light chain gene"
XX FT 1930..1935
XX FT /*tag= d
XX FT 1956..2104
XX FT /*tag= e
XX FT /note="kappa 3 flanking sequence"
XX FT 2105..4463
XX FT /*tag= f
XX FT /label=pBR322
XX FT /note="contg. ampicillinase gene and origin of
XX FT replication"
XX FT 4464..6971
XX FT /*tag= g
XX FT /label=psv2-gpt
XX FT /note="non-coding strand"
XX FT 6437..6547
XX FT /*tag= h
XX FT /label=cap_site
XX FT /note="from SV40(gpt)"
XX FT 6453..6469
XX FT /*tag= i
XX FT /label=origin region
XX FT /note="from SV40"
XX FT 6548..6619
XX FT /*tag= j
XX FT /rpt_type=direct
XX FT 6620..6691
XX FT /*tag= k
XX FT /rpt_type=direct
XX FT 6633..6650
XX FT /*tag= l
XX FT /note="from SV40"
XX FT 6716..6945
XX FT /*tag= m
XX FT /label=blocking element
XX FT /note="from lambda promoter"
XX FT 6946..6971
XX FT /*tag= n
XX FT /label=cap_site
XX FT /note="lambda mRNA"
XX FT 6972..7281
XX FT /*tag= o
XX FT /note="from Igh"
XX FT 7282..7466
XX FT /*tag= p
XX FT /note="from metallothioneine I"

```

```

protein_bind 7467..7533
/*tag= q
/label=cap_site
/note="from metallothioneine I"

WO8800242-A.
14-JAN-1988.
25-JUN-1987; 87WO-US01569.
26-JUN-1986; 86US-0879038.
(DAMO-) DAMON BIOTECH INC.
Gillies SD;
WPI; 1988-021592/03.
P-PSDB; AAP81913.
Prodn. of human tissue plasminogen activator - by myeloma cells,
pref. in presence of epsilon amino caproic acid.
Disclosure; Page ?; p; English.
The plasmid was prep'd by cleaving the very long 3' UT of t-PA cDNA
(which causes mRNA instability) 34 nucleotides downstream of stop
codon and inserting it into a pEMpl vector. The pEMpl vector was
constructed from the following components (see feature table):
(a) a 2.25 PvuII-BamHI fragment from pSV2-gpt contg. the SV40
enhancer and early region promoter, the E.coli gpt gene, the SV40
small tumour antigen intervening sequence, and the SV40 termination
and polyadenylation signals;
(b) a 2.3 kb PvuII-EcoRI fragment from pBR322 contg. the
ampicillinase gene and the bacterial origin of replication;
(c) a 0.3 kb PvuII-EcoRI fragment contg. an Ig heavy chain enhancer;
(d) a 0.25 kb SacI-BglII fragment contg. the metallothioneine I romoter;
(e) a 0.4 kb AvalI-HaeIII fragment from the 3' UT of Ig kappa light
chain gene; and
(f) a 0.26 kb XbaI-bstNI fragment contg. the Ig light chain promoter,
TAPAA sequence, and the initiation site.
The expression plasmid can be used to transform host cells esp.
J5581, ATCC CRL 9132.
SQ Sequence 7533 BP; 1918 A; 1892 C; 1847 G; 1875 T; 1 other;

Query Match 94.4%; Score 1065; DB 9; Length 7533;
Best Local Similarity 99.5%; Pred. No. 4.8e-225;
Matches 1068; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 56 CCCAGCGGCTCTGAGGAAACAGTCACTGCTACTTTGGGAATGGGTACGCTACCGTG 115
Db 633 CCCCTGCTGCTCTGAGGAAACAGTCACTGCTACTTTGGGAATGGGTACGCTACCGTG 692
Qy 116 GCACGACAGCCTCACCGAGTCGGGTGCTCTCTCCCTCCCGTGAATTCATGATCTGA 175
Db 693 GCACGACAGCCTCACCGAGTCGGGTGCTCTCTCCCTCCCGTGAATTCATGATCTGA 752
Qy 176 TAGGCAAGGTTTACACAGCAGAACCCAGTCGCCAGGCACTGGGCTGGGCAACATA 235
Db 753 TAGGCAAGGTTTACACAGCAGAACCCAGTCGCCAGGCACTGGGCTGGGCAACATA 812
Qy 236 ATTACTCCCGGAATCCTGATGGGATGCCAAGCCCTGGTGCACGTGCTGAAGAACCGCA 295
Db 813 ATTACTCCCGGAATCCTGATGGGATGCCAAGCCCTGGTGCACGTGCTGAAGAACCGCA 872
Qy 296 GGCTGACGTGGGAGTACTGTGATGTGCCCTCTCCCTCCAGCTGGGGCTGAGACAGTACA 355
Db 873 GGCTGACGTGGGAGTACTGTGATGTGCCCTCTCCCTCCAGCTGGGGCTGAGACAGTACA 932
Qy 356 GCCAGCCTCAGTTTCGCATCAAGAGAGGCTCTTTGGCGGACATCGCTCCACCCCTGGC 415
Db 933 GCCAGCCTCAGTTTCGCATCAAGAGAGGCTCTTTGGCGGACATCGCTCCACCCCTGGC 992

```

```
QY 416 AGGCTGCCATCTTTGCCAAGCACAGAGGTGCCCCGGAGAGCGGTTCTGTGCGGGGCA 475
DB 993 AGGCTGCCATCTTTGCCAAGCACAGAGGTGCCCCGGAGAGCGGTTCTGTGCGGGGCA 1052
QY 476 TACTCATCAGCTCCCTGCTGGATTCTCTCTGCGGCCCACTGCTCCAGGAGGTTTCGC 535
DB 1053 TACTCATCAGCTCCCTGCTGGATTCTCTCTGCGGCCCACTGCTCCAGGAGGTTTCGC 1112
QY 536 CCCACCACCTGACCGTGATCTTTGGCAGAAATACACCGGGTGTCTCCCTGGCAGGAGGAGC 595
DB 1113 CCCACCACCTGACCGTGATCTTTGGCAGAAATACACCGGGTGTCTCCCTGGCAGGAGGAGC 1172
QY 596 AGAAATTTGAAGTCGAAATAATACATTTGTCATAGAAATTCGATGATGACACTTACGACA 655
DB 1173 AGAAATTTGAAGTCGAAATAATACATTTGTCATAGAAATTCGATGATGACACTTACGACA 1232
QY 656 ATGACATTTGCGTGTCTGACGCTGAATCGGATTCGTCGCGCTGTGTCGCGAGGAGCAGC 715
DB 1233 ATGACATTTGCGTGTCTGACGCTGAATCGGATTCGTCGCGCTGTGTCGCGAGGAGCAGC 1292
QY 716 TGGTCCGCACTGTGTGCTTCCCCCGCGGACCTGACGCTCGGACCTGACGAGTG 775
DB 1293 TGGTCCGCACTGTGTGCTTCCCCCGCGGACCTGACGCTCGGACCTGACGAGTG 1352
QY 776 AGCTCTCCGGCTACCGCAAGCATAGGCTTGTCTCTCTTCTATTCGAGCGGCTGAAG 835
DB 1353 AGCTCTCCGGCTACCGCAAGCATAGGCTTGTCTCTCTTCTATTCGAGCGGCTGAAG 1412
QY 836 AGGCTCATGTGACACTGTACCCATCCAGCGCTGCACATCAACATTTACTTAAACAGAA 895
DB 1413 AGGCTCATGTGACACTGTACCCATCCAGCGCTGCACATCAACATTTACTTAAACAGAA 1472
QY 896 CAGTCACCGCAACATGTGTGTCTGGAGACACTCGGAGCGGGGCCCGCAAACT 955
DB 1473 CAGTCACCGCAACATGTGTGTCTGGAGACACTCGGAGCGGGGCCCGCAAACT 1532
QY 956 TGCACGACGCTGCCAGCGGATTCGGAGAGCCCTGCTGTGTGTGTGATGATGACGAGTGTCCGGGTG 1015
DB 1533 TGCACGACGCTGCCAGCGGATTCGGAGAGCCCTGCTGTGTGTGTGATGATGACGAGTGTCCGGGTG 1592
QY 1016 TGACTTTGCTGGGATCATCAGCTGGGCGCTGGGCTGTGGACAGAGATGTCCGGGTG 1075
DB 1593 TGACTTTGCTGGGATCATCAGCTGGGCGCTGGGCTGTGGACAGAGATGTCCGGGTG 1652
QY 1076 TGTACACAAAGTTACCACTACCTAGACTGGATTGCTGACAAACATGCGACCG 1128
DB 1653 TGTACACAAAGTTACCACTACCTAGACTGGATTGCTGACAAACATGCGACCG 1705
```

## RESULT 15

AAAT27588

ID AAAT27588 standard; DNA; 1680 BP.

XX AC AAAT27588;

XX DT 25-MAR-2003 (updated)

XX DT 06-AUG-1996 (first entry)

XX DE Novel plasminogen activator DNA.

XX Tissue plasminogen activator; tPA; alpha2-plasmin inhibitor;  
KW fibrinolytic; thrombolytic; fibrin; thrombosis; blood clotting;  
KW protein engineering; kringle; sa.  
XX Synthetic.

XX Key Location/Qualifiers  
FH sig\_peptide 1..105  
FT mat\_peptide 106..1677

FT /\*tag= a

FT /\*tag= b

US5504001-A.

XX 02-APR-1996.

XX 06-JUN-1994; 94US-0254485.

XX 25-NOV-1987; 87US-0125629.

XX 28-JAN-1992; 92US-082587.

XX 06-JUN-1994; 94US-0254485.

XX (ZYMO ) ZYMOGENETICS INC.

XX Foster DC;

XX WPI; 1996-187699/19.

XX P-PSDB; AAR96223.

XX Hybrid plasminogen activator comprises human tPA activator and

XX N-terminal crosslinking domain from alpha2-plasmin inhibitor -

XX useful to treat thrombosis and image blood clots

XX Example 3; Fig 11; 35pp; English.

XX A DNA construct (AAT27586) codes for a novel plasminogen

XX activator (AAR96222) in which the kringle K1 domain (AAR96221)

XX of plasminogen replaces the native K1 domain of tissue

XX plasminogen activator (tPA) (AAR96220). It was obt'd. by

XX inserting a synthetic plasminogen K1 domain DNA (AAT27586)

XX into tPA cDNA and can be expressed in transformed Escherichia

XX coli RRI cells (FERM P-9272). Novel plasminogen activators are

XX produced that show increased clot lysing specificity or plasma

XX half-life.

XX (Updated on 25-MAR-2003 to correct PF field.)

XX SQ Sequence 1680 BP; 387 A; 471 C; 479 G; 343 T; 0 other;

XX Query Match 94.3%; Score 1064; DB 17; Length 1680;

XX Best Local Similarity 99.5%; Pred. No. 6.6e-225;

XX Matches 1067; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 57 CCAGCGCGCTCTGAGGGAACAGTGACTGCTACTTTGGGAATGGCTCAGCTACCGTGG 116

DB 606 CCTGGAATGCTCTGAGGGAACAGTGACTGCTACTTTGGGAATGGCTCAGCTACCGTGG 665

QY 117 CACGACAGCTCACCAGTCCGGTGCCTCTGCTCCCGTGGAAATTCATGATCTCGAT 176

DB 666 CACGACAGCTCACCAGTCCGGTGCCTCTGCTCCCGTGGAAATTCATGATCTCGAT 725

QY 177 AGGCAAGTTTACACAGACAGAAACCCAGTCCCGGCACTGGGCTGGGCAACATAA 236

DB 726 AGGCAAGTTTACACAGACAGAAACCCAGTCCCGGCACTGGGCTGGGCAACATAA 785

QY 237 TTACTGCCGAATCCTGATGGGATGCCAAGCCCTGGTGCACGTCGTGGAAGAACCCAG 296

DB 786 TTACTGCCGAATCCTGATGGGATGCCAAGCCCTGGTGCACGTCGTGGAAGAACCCAG 845

QY 297 GCTGACGTGGAGTACTGTGATGTGCCCTCTGCTCCACCTGGGCTGGAGACGTACAG 356

DB 846 GCTGACGTGGAGTACTGTGATGTGCCCTCTGCTCCACCTGGGCTGGAGACGTACAG 905

QY 357 CCAGCCTCAGTTTCCGATCAAGAGAGGGCTCTTCCCGGACATCGCTCCACCCCTGGCA 416

DB 906 CCAGCCTCAGTTTCCGATCAAGAGAGGGCTCTTCCCGGACATCGCTCCACCCCTGGCA 965

QY 417 GGCTGCCATCTTTGCCAAGCACAGGAGGTGCCCGGAGAGCGGTTCTGTGCGGGGCGAT 476

DB 966 GGCTGCCATCTTTGCCAAGCACAGGAGGTGCCCGGAGAGCGGTTCTGTGCGGGGCGAT 1025

QY 477 ACTCATCAGCTCTGCTGGATTCTCTGCTGCCGCCCACTGCTCCAGGAGAGGTTTCGCC 536

DB 1026 ACTCATCAGCTCTGCTGGATTCTCTGCTGCCGCCCACTGCTCCAGGAGAGGTTTCGCC 1085

QY 537 CCACCACCTGACGGTGATCTTTGGGCGAGAAATATACCGGGTGGTCCCTGGCGAGGAGCA 596

Db 1086 CCACCACTGACGGTGAATCTTGGGCAGAAACATACCGGTGGTCCCTGGCGAGGAGGCA 1145  
QY 597 GAAATTTGAAGTCGAAAATACATTTGCCATAAGAAATTCGATGATGACACTTACGACAA 656  
Db 1146 GAAATTTGAAGTCGAAAATACATTTGCCATAAGAAATTCGATGATGACACTTACGACAA 1205  
QY 657 TGACATTCGGCTGCTGCAGCTGAAATCGAATTCGTTCCCGCTGTGCCAGGAGAGCAGCGT 716  
Db 1206 TGACATTCGGCTGCTGCAGCTGAAATCGAATTCGTTCCCGCTGTGCCAGGAGAGCAGCGT 1265  
QY 717 GGTCCGCACTGTGTGCTTCCCGCGGAGCACTGCAGCTGCCGACTGGACGGAGTGTGA 776  
Db 1266 GGTCCGCACTGTGTGCTTCCCGCGGAGCACTGCAGCTGCCGACTGGACGGAGTGTGA 1325  
QY 777 GCTCTCCGGCTACGGCAAGCATGAGGCTTGTCTCCTTTCTATTTCGGAGCGGCTGAAGGA 836  
Db 1326 GCTCTCCGGCTACGGCAAGCATGAGGCTTGTCTCCTTTCTATTTCGGAGCGGCTGAAGGA 1385  
QY 837 GGCTCATGTCAAGCTGTACCCATCCAGCGGCTGCACATCACAACATTTTACTTAAACAGAAC 896  
Db 1386 GGCTCATGTCAAGCTGTACCCATCCAGCGGCTGCACATCACAACATTTTACTTAAACAGAAC 1445  
QY 897 AGTCAACCGACAACATGTGTGTGTGGAGACACTCGGAGCGCGGCCCCCAGGCAAACTT 956  
Db 1446 AGTCAACCGACAACATGTGTGTGTGGAGACACTCGGAGCGCGGCCCCCAGGCAAACTT 1505  
QY 957 GCACGACGCTGCAGGGCGGATTCGGGAGGCCCCCTGTGTGTCTGAACGATGGCCGAT 1016  
Db 1506 GCACGACGCTGCAGGGCGGATTCGGGAGGCCCCCTGTGTGTCTGAACGATGGCCGAT 1565  
QY 1017 GACTTTGGTGGGCATCATCAGCTGGGCGCTGGGCTGTGCAGACAGAGGATGTCCCGGGTGT 1076  
Db 1566 GACTTTGGTGGGCATCATCAGCTGGGCGCTGGGCTGTGCAGACAGAGGATGTCCCGGGTGT 1625  
QY 1077 GTACACAAAGGTTACCAACTACCTAGACTGGATTCGTGACAAACATGCCGACCG 1128  
Db 1626 GTACACAAAGGTTACCAACTACCTAGACTGGATTCGTGACAAACATGCCGACCG 1677

Search completed: August 18, 2003, 23:28:52  
Job time : 398.386 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 18, 2003, 22:51:23 ; Search time 88.2804 Seconds  
(without alignments)  
5639.757 Million cell updates/sec

Title: US-09-987-455-5

Perfect score: 1128

Sequence: 1 atgaaaagacagctatcgc.....ttcgtgacaacatgcgacg 1128

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA.\*

1: /cgn2\_6/ptodata/2/ina/5A\_COMB.seq:\*  
2: /cgn2\_6/ptodata/2/ina/5B\_COMB.seq:\*  
3: /cgn2\_6/ptodata/2/ina/6A\_COMB.seq:\*  
4: /cgn2\_6/ptodata/2/ina/6B\_COMB.seq:\*  
5: /cgn2\_6/ptodata/2/ina/PCTUS\_COMB.seq:\*  
6: /cgn2\_6/ptodata/2/ina/backfiles.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1065	94.4	1419	2	US-08-811-949-62
2	1065	94.4	1848	3	US-08-814-412-10
3	1065	94.4	2101	2	US-08-811-949-42
4	1065	94.4	2162	1	US-08-119-512-3
5	1065	94.4	2162	1	US-08-488-015B-3
6	1065	94.4	2162	1	US-08-488-015B-25
7	1063.4	94.3	1314	2	US-08-811-949-48
8	1063.4	94.3	1738	6	5200340-1
9	1063.4	94.3	1957	2	US-08-883-795A-39
10	1063.4	94.3	2457	6	5344773-1
11	1063.4	94.3	7360	1	US-08-286-740-1
12	1063.4	94.3	7360	5	PCT-US95-09576-1
13	1062	94.1	1068	2	US-08-811-949-44
14	1061.8	94.1	1314	2	US-08-811-949-50
15	1061.8	94.1	1974	2	US-08-811-949-38
16	1060.4	94.0	1170	2	US-08-811-949-64
17	1058.8	93.9	1068	2	US-08-811-949-46
18	1057	93.7	1314	2	US-08-811-949-54
19	1056.4	93.7	1068	1	US-08-137-116-2
20	1054.8	93.5	1065	1	US-08-427-640-1
21	1054.8	93.5	1137	4	US-09-553-498-9
22	1054.8	93.5	1137	4	US-09-618-869-9
23	1054	93.4	1068	2	US-08-811-949-52
24	1053.8	93.4	1314	2	US-08-811-949-56
25	1052.4	93.3	1170	2	US-08-811-949-66
26	1050	93.1	1068	1	5223256-3
27	1046.8	92.8	1065	1	US-08-427-640-5

28 1038 92.0 1068 2 US-08-811-949-58 Sequence 58, Appl  
29 1035.4 91.8 1068 1 US-08-427-640-3 Sequence 3, Appl  
30 1030 91.3 1065 2 US-08-811-949-60 Sequence 60, Appl  
31 953.4 84.5 2497 6 5185259-2 Patent No. 5185259  
32 924 81.9 1068 1 US-08-427-640-7 Sequence 7, Appl  
33 798 70.7 1163 2 US-08-558-269-5 Sequence 5, Appl  
34 798 70.7 1163 3 US-09-410-882-5 Sequence 5, Appl  
35 625.8 55.5 1727 6 5244676-4 Patent No. 5244676  
36 625.6 55.5 1605 6 5244676-1 Patent No. 5244676  
37 472 41.8 472 2 US-08-811-949-40 Sequence 40, Appl  
38 414 36.7 453 6 5200340-3 Patent No. 5200340  
39 267.4 23.7 1724 6 5200340-5 Patent No. 5200340  
40 170.8 15.1 1236 1 US-07-957-039A-7 Sequence 7, Appl  
41 170.8 15.1 1475 4 US-09-643-597-122 Sequence 122, App  
42 170.8 15.1 1475 4 US-09-480-884A-122 Sequence 122, App  
43 170.8 15.1 1475 4 US-09-542-615A-122 Sequence 122, App  
44 170.8 15.1 1475 4 US-09-606-421B-122 Sequence 122, App  
45 170.8 15.1 2301 6 5188829-2 Patent No. 5188829

## ALIGNMENTS

RESULT 1  
US-08-811-949-62  
; Sequence 62, Application US/08811949  
; Patent No. 5840533  
; GENERAL INFORMATION:  
; APPLICANT: NIWA, MINEO  
; APPLICANT: SAITO, YOSHIMASA  
; APPLICANT: SASAKI, HITOSHI  
; APPLICANT: HAYASHI, MASAKO  
; APPLICANT: NOTANI, JOUJI  
; APPLICANT: KOBAYASHI, MASAKAZU  
; TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR  
; NUMBER OF SEQUENCES: 67  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
; ADDRESSEE: P.C.  
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400  
; CITY: ARLINGTON  
; STATE: VA  
; COUNTRY: USA  
; ZIP: 22202  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/811,949  
; FILING DATE: 05-MAR-1997  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: OBLON, NORMAN F.  
; REGISTRATION NUMBER: 24,618  
; REFERENCE/DOCKET NUMBER: 18-966-0  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 703-413-3000  
; TELEFAX: 703-413-2220  
; INFORMATION FOR SEQ ID NO: 62:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1419 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: circular  
; MOLECULE TYPE: DNA (genomic)  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 1..1416  
US-08-811-949-62

Query Match 94.4%; Score 1065; DB 2; Length 1419;



QY 416 AGGCTGCCATCTTTCCCAAGCAGAGAGGTGCGCCGAGAGGGTTCTCTGCGGGGCA 475  
Db 1045 AGGCTGCCATCTTTCCCAAGCAGAGAGGTGCGCCGAGAGGGTTCTCTGCGGGGCA 1104  
QY 476 TACTCATCAGCTCTGCTGGATTCTCTGCGGCCACTGCTTCCAGAGAGGTTTCGCG 535  
Db 1105 TACTCATCAGCTCTGCTGGATTCTCTGCGGCCACTGCTTCCAGAGAGGTTTCGCG 1164  
QY 536 CCCACCACCTGACGGTGATCTTTGGCAGAACATACCGGGTGCTCTCCGCGAGAGGAGC 595  
Db 1165 CCCACCACCTGACGGTGATCTTTGGCAGAACATACCGGGTGCTCTCCGCGAGAGGAGC 1224  
QY 596 AGAAATTTGAAGTCGAAAATACATGTCCTAAGGAATTCGATGATGACACTTACGCA 655  
Db 1225 AGAAATTTGAAGTCGAAAATACATGTCCTAAGGAATTCGATGATGACACTTACGCA 1284  
QY 656 ATGACATTCGCTGCTGAGCTGAATTCGCTCCGCTGTCGCCAGAGAGCAGCG 715  
Db 1285 ATGACATTCGCTGCTGAGCTGAATTCGCTCCGCTGTCGCCAGAGAGCAGCG 1344  
QY 716 TGGTCCGCACTGTGTGCTTCCCGCGGACCTGACGCTGCGGACTGGACGGAGTG 775  
Db 1345 TGGTCCGCACTGTGTGCTTCCCGCGGACCTGACGCTGCGGACTGGACGGAGTG 1404  
QY 776 AGCTCTCCGCTACGCGAAGCATGAGGCTTGTCTCTTTCTATTTCGAGCGGCTGAAG 835  
Db 1405 AGCTCTCCGCTACGCGAAGCATGAGGCTTGTCTCTTTCTATTTCGAGCGGCTGAAG 1464  
QY 836 AGGCTCAGTCAGACTGTACCCATCCAGCGCTGCACATCAACATTTACTTAACAGAA 895  
Db 1465 AGGCTCAGTCAGACTGTACCCATCCAGCGCTGCACATCAACATTTACTTAACAGAA 1524  
QY 896 CAGTCACCGACAACTGTGTGTGTGAGACACTCGGAGCGGGGCCCGAGCAAACT 955  
Db 1525 CAGTCACCGACAACTGTGTGTGTGAGACACTCGGAGCGGGGCCCGAGCAAACT 1584  
QY 956 TGCAGACGCTGCCAGGGGATTCGGAGGCGCCCTGCTGTGTCTGAACGATGCGCGCA 1015  
Db 1585 TGCAGACGCTGCCAGGGGATTCGGAGGCGCCCTGCTGTGTCTGAACGATGCGCGCA 1644  
QY 1016 TGACTTTGCTGGCATCATCAGCTGGGCGCTGGGCTGTGGACAGAGATGTCGCGGGTG 1075  
Db 1645 TGACTTTGCTGGCATCATCAGCTGGGCGCTGGGCTGTGGACAGAGATGTCGCGGGTG 1704  
QY 1076 TGTACAAAAGGTTACCAACTACCTAGACTGGATTGCTGACAAACATGCGACCG 1128  
Db 1705 TGTACAAAAGGTTACCAACTACCTAGACTGGATTGCTGACAAATGCGACCG 1757

RESULT 3

US-08-811-949-42  
; Sequence 42, Application US/08811949  
; Patent No. 5840533  
; GENERAL INFORMATION:  
; APPLICANT: NIWA, MINEO  
; APPLICANT: SAITO, YOSHIMASA  
; APPLICANT: SASAKI, HITOSHI  
; APPLICANT: HAYASHI, MASAKO  
; APPLICANT: NOTANI, JOUJI  
; APPLICANT: KOBAYASHI, MASAKAZU  
; TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR  
; NUMBER OF SEQUENCES: 67  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: P.C.  
; ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,  
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400  
; CITY: ARLINGTON  
; STATE: VA  
; COUNTRY: USA  
; ZIP: 22202  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/811,949  
; FILING DATE: 05-MAR-1997  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: OBLON, NORMAN F.  
; REGISTRATION NUMBER: 24,618  
; REFERENCE/DOCKET NUMBER: 18-966-0  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 703-413-3000  
; TELEFAX: 703-413-2220  
; INFORMATION FOR SEQ ID NO: 42:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2101 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: circular  
; MOLECULE TYPE: DNA (genomic)  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 25...1710  
; US-08-811-949-42

Query Match 94.4%; Score 1065; DB 2; Length 2101;  
Best Local Similarity 99.5%; Pred. No. 9.1e-290;  
Matches 1068; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 56 CCCAGGCGGCTCTGAGGGAAACAGTCACTGCTACTTTGGGAATGGGTACGCTTACCGTG 115  
Db 638 CCCCTGCTGCTCTGAGGGAAACAGTCACTACTTTTGGGAATGGGTACGCTTACCGTG 697  
QY 116 GCACGACAGCTCACCGAGTCCGCTGCTTCCCTCCGCTGGAATTCATGATCTCTGA 175  
Db 698 GCACGACAGCTCACCGAGTCCGCTGCTTCCCTCCGCTGGAATTCATGATCTCTGA 757  
QY 176 TAGGCAAGGTTTACAGACAGAACCCAGTCCAGGCACTGGGCTGGGCAACATA 235  
Db 758 TAGGCAAGGTTTACAGACAGAACCCAGTCCAGGCACTGGGCTGGGCAACATA 817  
QY 236 ATACTGCGGATCTCTGATGGGATGCCAAGCCCTGCTGTCAGTCTGAGAACCGCA 295  
Db 818 ATACTGCGGATCTCTGATGGGATGCCAAGCCCTGCTGTCAGTCTGAGAACCGCA 877  
QY 296 GGCTGAGTGGGAGTCTGTGATGCTGCTCCCTGCTCCACCTGCGGCTGAGACAGTACA 355  
Db 878 GGCTGAGTGGGAGTCTGTGATGCTGCTCCCTGCTCCACCTGCGGCTGAGACAGTACA 937  
QY 356 GCCAGCCTCAGTTTTCGATCAAGAGGAGGCTCTTTGCGCGACATCGCCTCCACCCCTGGC 415  
Db 938 GCCAGCCTCAGTTTTCGATCAAGAGGAGGCTCTTTGCGCGACATCGCCTCCACCCCTGGC 997  
QY 416 AGGCTGCCATCTTTGCCAAGCAGAGAGGTGCGCCGAGAGCGGTTCTCTGTCGGGGGCA 475  
Db 998 AGGCTGCCATCTTTGCCAAGCAGAGAGGTGCGCCGAGAGCGGTTCTCTGTCGGGGGCA 1057  
QY 476 TACTCATCAGCTCTGCTGGATTCTCTGCGGCCACTGCTTCCAGAGAGGTTTCGCG 535  
Db 1058 TACTCATCAGCTCTGCTGGATTCTCTGCGGCCACTGCTTCCAGAGAGGTTTCGCG 1117  
QY 536 CCCACCACCTGACGGTGATCTTTGGCAGAACATACCGGGTGCTCTCCGCGAGAGGAGC 595  
Db 1118 CCCACCACCTGACGGTGATCTTTGGCAGAACATACCGGGTGCTCTCCGCGAGAGGAGC 1177  
QY 596 AGAAATTTGAAGTCGAAAATACATGTCCTAAGGAATTCGATGATGACACTTACGCA 655  
Db 1178 AGAAATTTGAAGTCGAAAATACATGTCCTAAGGAATTCGATGATGACACTTACGCA 1237  
QY 656 ATGACATTCGCTGCTGAGCTGAATTCGATGCTCCGCTGTCGCCAGAGAGCAGCG 715  
Db 1238 ATGACATTCGCTGCTGAGCTGAATTCGATGCTCCGCTGTCGCCAGAGAGCAGCG 1297

```

QY 716 TGGTCCGCACTGTGTGCTTCCCGCGGACCTGACGTGCGGACCTGGAGGAGTGTG 775
Db 1298 TGGTCCGCACTGTGTGCTTCCCGCGGACCTGACGTGCGGACCTGGAGGAGTGTG 1357
QY 776 AGCTCTCGGCTACCGCAAGCATGAGGCTTGTCTCTTTCTATTTCGAGGCGGTGAAGG 835
Db 1358 AGCTCTCGGCTACCGCAAGCATGAGGCTTGTCTCTTTCTATTTCGAGGCGGTGAAGG 1417
QY 836 AGGCTCATGTGACAGTGTACCATCCAGCGCTGCACATCATCAACATTTACTTAACAGAA 895
Db 1418 AGGCTCATGTGACAGTGTACCATCCAGCGCTGCACATCATCAACATTTACTTAACAGAA 1477
QY 896 CAGTCACCGACAACTGTGTGTGTGTGAGACACTCGGAGCGGCGGCGCCAGGCAAACT 955
Db 1478 CAGTCACCGACAACTGTGTGTGTGTGAGACACTCGGAGCGGCGGCGCCAGGCAAACT 1537
QY 956 TGCACGACGCTGCCAGGCGGATTCGGAGGCGCCCTGCTGTGTGTGTGAACGATGCGCGCA 1015
Db 1538 TGCACGACGCTGCCAGGCGGATTCGGAGGCGCCCTGCTGTGTGTGTGAACGATGCGCGCA 1597
QY 1016 TGACTTTGTGGGCACTCATGCTGCGGCGCTGGGCTGTGGACAGAGGATGTCCGGGTG 1075
Db 1598 TGACTTTGTGGGCACTCATGCTGCGGCGCTGGGCTGTGGACAGAGGATGTCCGGGTG 1657
QY 1076 TGTACACAAAGGTTACCACTACCTAGACTGTGATTCGTGACAACTGCGACCG 1128
Db 1658 TGTACACAAAGGTTACCACTACCTAGACTGTGATTCGTGACAACTGCGACCG 1710

```

## RESULT 4

```

US-08-119-512-3
; Sequence 3, Application US/08119512
; Patent No. 5498531
; GENERAL INFORMATION:
; APPLICANT: Jarrell, Kevin A.
; TITLE OF INVENTION: INTRON MEDIATED RECOMBINANT TECHNIQUES
; TITLE OF INVENTION: AND REAGENTS
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII (text)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/119,512
; FILING DATE: 10-SEP-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: HUI-008
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2162 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: both
; MOLECULE TYPE: other nucleic acid
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 82..334
; OTHER INFORMATION: /product= "Signal Sequence and
; OTHER INFORMATION: Finger-like domain"

```

```

; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 335..447
; OTHER INFORMATION: /product= "EGF-like domain"
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 448..714
; OTHER INFORMATION: /product= "Kringle-1 domain"
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 715..972
; OTHER INFORMATION: /product= "Kringle-2 domain"
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 973..2162
; OTHER INFORMATION: /product= "Catalytic domain"
; US-08-119-512-3
Query Match 94.4%; Score 1065; DB 1; Length 2162;
Best Local Similarity 99.5%; Pred. No. 9.2e-290;
Matches 1068; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 56 CCCAGGCGGCTCTGTAGGGAACAGTGTACTTTTGGGAATGGGTGACGCTTACCGTG 115
Db 695 CCCCTGCTCTGTAGGGAACAGTGTACTTTTGGGAATGGGTGACGCTTACCGTG 754
QY 116 GCACGACAGCTCACCGAGTGGGTGCTCTGCTCCGCTCCGCTGGAATTCATGATCCTGA 175
Db 755 GCACGACAGCTCACCGAGTGGGTGCTCTGCTCCGCTCCGCTGGAATTCATGATCCTGA 814
QY 176 TAGCAAGGTTTACACAGCACAGAACCCAGTGCACAGGCACTGGGCTTGGGCAACATA 235
Db 815 TAGCAAGGTTTACACAGCACAGAACCCAGTGCACAGGCACTGGGCTTGGGCAACATA 874
QY 236 ATTACTCCGGAATCTGTATGGGATGCCAAGCCCTGGTGCCACGTCTGCTGAAGAACCGCA 295
Db 875 ATTACTCCGGAATCTGTATGGGATGCCAAGCCCTGGTGCCACGTCTGCTGAAGAACCGCA 934
QY 296 GGCTGAGTGGGAGTACTGTGCTCCCTGCTGCTCCAGCTGGGCTGAGACAGTACA 355
Db 935 GGCTGAGTGGGAGTACTGTGCTCCCTGCTGCTCCAGCTGGGCTGAGACAGTACA 994
QY 356 GCCAGCCTCAGTTTTCGATCAAAAGGAGGCTCTTTGCGCCGACATGCGCTCCACCCCTGGC 415
Db 995 GCCAGCCTCAGTTTTCGATCAAAAGGAGGCTCTTTGCGCCGACATGCGCTCCACCCCTGGC 1054
QY 416 AGGCTGCCATCTTTGCCAAGCAGAGGCTGCGCCGAGAGAGCGGTCTCTGTGCGGGGCA 475
Db 1055 AGGCTGCCATCTTTGCCAAGCAGAGGCTGCGCCGAGAGAGCGGTCTCTGTGCGGGGCA 1114
QY 476 TACTCATCAGCTCTGCTGATTTCTCTGCGCGCCCACTGCTTCCAGGAGGTTTCCGC 535
Db 1115 TACTCATCAGCTCTGCTGATTTCTCTGCGCGCCCACTGCTTCCAGGAGGTTTCCGC 1174
QY 536 CCCACCACTGAGGCTGATCTTGGGAGAAACATACCGGGTGGTCCCTTGGGAGAGGAGC 595
Db 1175 CCCACCACTGAGGCTGATCTTGGGAGAAACATACCGGGTGGTCCCTTGGGAGAGGAGC 1234
QY 596 AGAAATTTGAAGTCGAAAAATAATTTGTCATTAAGGAATTCGATGATGACACTTACGACA 655
Db 1235 AGAAATTTGAAGTCGAAAAATAATTTGTCATTAAGGAATTCGATGATGACACTTACGACA 1294
QY 656 ATGACATTGCGCTGCTGACGTGAAATCGATTCGTCGCGCCCTGTCAGGAGGAGCAGCG 715
Db 1295 ATGACATTGCGCTGCTGACGTGAAATCGATTCGTCGCGCCCTGTCAGGAGGAGCAGCG 1354
QY 716 TGGTCCGCACTGTGTGCTTCCCGCGGACCTGACGTGCGGCACTGACGAGGTGTG 775
Db 1355 TGGTCCGCACTGTGTGCTTCCCGCGGACCTGACGTGCGGCACTGACGAGGTGTG 1414
QY 776 AGCTCTCCGCTACCGCAAGCATGAGGCTTGTCTCTTTCTATTTCGAGGCGGTGAAGG 835
Db 1415 AGCTCTCCGCTACCGCAAGCATGAGGCTTGTCTCTTTCTATTTCGAGGCGGTGAAGG 1474

```



Db 1535 CAGTCACCGACAAACATGCTGTGTCTGTGAGACACTCGGAGCGCGGCCCCAGGCAAACT 1594  
Qy 956 TGCACGAGCGCTGCCAGGCGGATTCGGAGGCGCCCTCGTGTCTGAACGATGCGCGCA 1015  
Db 1595 TGCACGAGCGCTGCCAGGCGGATTCGGAGGCGCCCTCGTGTCTGAACGATGCGCGCA 1654  
Qy 1016 TGACTTTGGTGGGCATCATCAGCTGGGCGCTGGGCTGTGGACAGAGGATGTCCGGGTG 1075  
Db 1655 TGACTTTGGTGGGCATCATCAGCTGGGCGCTGGGCTGTGGACAGAGGATGTCCGGGTG 1714  
Qy 1076 TGTACAAAAGTTTACCAACTACCTAGACTGGATTCTGTGACAAATGCGACCG 1128  
Db 1715 TGTACAAAAGTTTACCAACTACCTAGACTGGATTCTGTGACAAATGCGACCG 1767

## RESULT 6

US-08-488-015B-25  
; Sequence 25, Application US/08488015B  
; Patent No. 5780272

## GENERAL INFORMATION:

; APPLICANT: Jarrell, Kevin A.  
; TITLE OF INVENTION: INTRON-MEDIATED RECOMBINANT TECHNIQUES

; TITLE OF INVENTION: AND REAGENTS

; NUMBER OF SEQUENCES: 29

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Foley, Hoag & Eliot

; STREET: One Post Office Square

; CITY: Boston

; STATE: MA

; COUNTRY: USA

; ZIP: 02109

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: ASCII (text)

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/488,015B

; FILING DATE: 07-JUN-1995

; ATTORNEY/AGENT INFORMATION:

; NAME: Vincent, Matthew P.

; REGISTRATION NUMBER: 36,709

; REFERENCE/DOCKET NUMBER: HUV-008.02

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (617) 832-1000

; TELEFAX: (617) 832-7000

; INFORMATION FOR SEQ ID NO: 25:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 2162 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: cDNA

US-08-488-015B-25

Query Match 94.4%; Score 1065; DB 1; Length 2162;

Best Local Similarity 99.5%; Pred. No. 9.2e-290;

Matches 1068; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 56 CCCAGCGCGCTCTGAGGAAACAGTGACTGTCTTTGGGAATGGGTGAGCTACCGTG 115  
Db 695 CCCCTGCTCTCTGAGGAAACAGTGACTGTCTTTGGGAATGGGTGAGCTACCGTG 754  
Qy 116 GCACGACAGCTTACCGAGTGGGTGCTCTGCTCCGCTGGGAATTCATGATCCTGA 175  
Db 755 GCACGACAGCTTACCGAGTGGGTGCTCTGCTCCGCTGGGAATTCATGATCCTGA 814  
Qy 176 TAGGCAAGTTTACACAGCAGAAACCCAGTGCCAGGCACTGGGCTGGGCAACATA 235  
Db 815 TAGGCAAGTTTACACAGCAGAAACCCAGTGCCAGGCACTGGGCTGGGCAACATA 874  
Qy 236 ATTACTCGCGGAATCCTGATGGGATGCCAAGCCCTGGTGCACTGTGAGAACCGCA 295

Db 875 ATTACTCGCGGAATCCTGATGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCA 934  
Qy 296 GGCTGAGCTGGGAGTACTGTGATGTGCCCTCTCTCCACCTGGCGCTGAGACAGTACA 355  
Db 935 GGCTGAGCTGGGAGTACTGTGATGTGCCCTCTCTCCACCTGGCGCTGAGACAGTACA 994  
Qy 356 GCCACGCTCAGTTTCGCATCAAAGAGGGCTCTTTCGCCGACATGCCCTCCACCCCTGGC 415  
Db 995 GCCACGCTCAGTTTCGCATCAAAGAGGGCTCTTTCGCCGACATGCCCTCCACCCCTGGC 1054  
Qy 416 AGGCTGCCATCTTTTGCACAGCAGAGAGTGCCTCCGAGAGCGGTTCCTGTGCGGGGCA 475  
Db 1055 AGGCTGCCATCTTTTGCACAGCAGAGAGTGCCTCCGAGAGCGGTTCCTGTGCGGGGCA 1114  
Qy 476 TACTCATCAGCTCCTGTGGATTCTCTGCGCGCCACTGCTCCAGGAGGTTTCCGC 535  
Db 1115 TACTCATCAGCTCCTGTGGATTCTCTGCGCGCCACTGCTCCAGGAGGTTTCCGC 1174  
Qy 536 CCCACCACTGACGGTGATCTTGGGCGAGAAATACCGGGTGGTCCCTGGCGAGGAGGAGC 595  
Db 1175 CCCACCACTGACGGTGATCTTGGGCGAGAAATACCGGGTGGTCCCTGGCGAGGAGGAGC 1234  
Qy 596 AGAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA 655  
Db 1235 AGAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA 1294  
Qy 656 ATGACATTTGCGCTGCTGAGTGAATCGGATTCGCTCCGCTGTCGCCAGGAGAGCAGCG 715  
Db 1295 ATGACATTTGCGCTGCTGAGTGAATCGGATTCGCTCCGCTGTCGCCAGGAGAGCAGCG 1354  
Qy 716 TGGTCCGCACTGTGTGCTTCCCGCGGACCTGACAGCTGCCGAGCTGGAACGAGTGTG 775  
Db 1355 TGGTCCGCACTGTGTGCTTCCCGCGGACCTGACAGCTGCCGAGCTGGAACGAGTGTG 1414  
Qy 776 AGCTCTCGGCTACGCAAGCATGAGCGTGTCTCTTCTATTCGAGCGGCTGAAGG 835  
Db 1415 AGCTCTCGGCTACGCAAGCATGAGCGTGTCTCTTCTATTCGAGCGGCTGAAGG 1474  
Qy 836 AGGCTCATGTGACAGTGTACCCATCCAGCGCTGCATCACACATTTACTTAACAGAA 895  
Db 1475 AGGCTCATGTGACAGTGTACCCATCCAGCGCTGCATCACACATTTACTTAACAGAA 1534  
Qy 896 CAGTCACCGACAAACATGCTGTGTCTGGAGACACTCGGAGCGCGGCGCCAGGCAAACT 955  
Db 1535 CAGTCACCGACAAACATGCTGTGTCTGGAGACACTCGGAGCGCGGCGCCAGGCAAACT 1594  
Qy 956 TGACAGACGCTGCGAGGCGATTCGGAGGCGCCCTGGTGTGTCTGAACGATGCGCGCA 1015  
Db 1595 TGACAGACGCTGCGAGGCGATTCGGAGGCGCCCTGGTGTGTCTGAACGATGCGCGCA 1654  
Qy 1016 TGACTTTGGTGGGCATCATCAGCTGGGCGCTGGGCTGGGAGAGGATGTCGCGGGTG 1075  
Db 1655 TGACTTTGGTGGGCATCATCAGCTGGGCGCTGGGCTGGGAGAGGATGTCGCGGGTG 1714  
Qy 1076 TGATACAAAAGTTTACCAACTACCTAGACTGGATTCTGTGACAAATGCGACCG 1128  
Db 1715 TGATACAAAAGTTTACCAACTACCTAGACTGGATTCTGTGACAAATGCGACCG 1767

## RESULT 7

US-08-811-949-48

; Sequence 48, Application US/08811949

; Patent No. 5840533

; GENERAL INFORMATION:

; APPLICANT: NIWA, MINEO

; APPLICANT: SAITO, YOSHIMASA

; APPLICANT: SASAKI, HITOSHI

; APPLICANT: HAYASHI, MASAKO

; APPLICANT: NOTANI, JOUJI

; APPLICANT: KOBAYASHI, MASAKAZU

; TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR

; NUMBER OF SEQUENCES: 67

; CORRESPONDENCE ADDRESS:

0:	1:	2:	3:	4:	5:	6:	7:	8:	9:
3 115	3 116	3 117	3 118	3 119	3 120	3 121	3 122	3 123	3 124
3 298	3 299	3 300	3 301	3 302	3 303	3 304	3 305	3 306	3 307
A 175	A 176	A 177	A 178	A 179	A 180	A 181	A 182	A 183	A 184
A 358	A 359	A 360	A 361	A 362	A 363	A 364	A 365	A 366	A 367
A 235	A 236	A 237	A 238	A 239	A 240	A 241	A 242	A 243	A 244
A 418	A 419	A 420	A 421	A 422	A 423	A 424	A 425	A 426	A 427
A 295	A 296	A 297	A 298	A 299	A 300	A 301	A 302	A 303	A 304
A 478	A 479	A 480	A 481	A 482	A 483	A 484	A 485	A 486	A 487
A 355	A 356	A 357	A 358	A 359	A 360	A 361	A 362	A 363	A 364
A 538	A 539	A 540	A 541	A 542	A 543	A 544	A 545	A 546	A 547
C 415	C 416	C 417	C 418	C 419	C 420	C 421	C 422	C 423	C 424
C 598	C 599	C 600	C 601	C 602	C 603	C 604	C 605	C 606	C 607
A 475	A 476	A 477	A 478	A 479	A 480	A 481	A 482	A 483	A 484
A 658	A 659	A 660	A 661	A 662	A 663	A 664	A 665	A 666	A 667
C 535	C 536	C 537	C 538	C 539	C 540	C 541	C 542	C 543	C 544
C 718	C 719	C 720	C 721	C 722	C 723	C 724	C 725	C 726	C 727
C 595	C 596	C 597	C 598	C 599	C 600	C 601	C 602	C 603	C 604

[illegible]

```
QY 296 GGCTGACGTGGAGTACTGTGATGTCCTCTGCTCCACCTGCGGCTGAGACAGTACA 355
DB 869 GGCTGACGTGGAGTACTGTGATGTCCTCTGCTCCACCTGCGGCTGAGACAGTACA 928
QY 356 GCCAGCCTCAGTTTCGATCAAAAGAGAGGCTCTTCCGCGACATCGCTCCACCCCTGGC 415
DB 929 GCCAGCCTCAGTTTCGATCAAAAGAGAGGCTCTTCCGCGACATCGCTCCACCCCTGGC 988
QY 416 AGGCTGCCATCTTTGCCAAGCACAGGAGGTGCGCGGAGAGGGTTCTGTGCGGGGCA 475
DB 989 AGGCTGCCATCTTTGCCAAGCACAGGAGGTGCGCGGAGAGGGTTCTGTGCGGGGCA 1048
QY 476 TACTCATCAGCTCTGCTGGAATCTCTGCGCGCCACTGCTTCCAGGAGAGTTTCCGC 535
DB 1049 TACTCATCAGCTCTGCTGGAATCTCTGCGCGCCACTGCTTCCAGGAGAGTTTCCGC 1108
QY 536 CCCACCACCTGACGGTGTCTTGGGAGACATACCGGGTGTCTTCCGCGAGGAGGAGC 595
DB 1109 CCCACCACCTGACGGTGTCTTGGGAGACATACCGGGTGTCTTCCGCGAGGAGGAGC 1168
QY 596 AGAAATTTGAAGTCGAAATATACATTGTCATAGGAATTCGATGACACTTACGACA 655
DB 1169 AGAAATTTGAAGTCGAAATATACATTGTCATAGGAATTCGATGACACTTACGACA 1228
QY 656 ATGACATTTGCGTCTGCGAGCTGAAATCGGATTCGTCGCGCTGTGCCAGGAGAGCAGC 715
DB 1229 ATGACATTTGCGTCTGCGAGCTGAAATCGGATTCGTCGCGCTGTGCCAGGAGAGCAGC 1288
QY 716 TGGTCCGCACTGTGTGCTTCCCGCGGAGACCTGCGAGCTGCGGAGTGGACGGATGTG 775
DB 1289 TGGTCCGCACTGTGTGCTTCCCGCGGAGACCTGCGAGCTGCGGAGTGGACGGATGTG 1348
QY 776 AGCTCTCGGCTACGGCAAGCATGAGGCTTGTCTCTTCTTCTATTCGAGGCGGTGAAG 835
DB 1349 AGCTCTCGGCTACGGCAAGCATGAGGCTTGTCTCTTCTTCTATTCGAGGCGGTGAAG 1408
QY 836 AGGCTCATGTGAGTGTACCATCCAGCGCTGCATACATCAACATTTACTTAACAGAA 895
DB 1409 AGGCTCATGTGAGTGTACCATCCAGCGCTGCATACATCAACATTTACTTAACAGAA 1468
QY 896 CAGTCACCGACAACATGCTGTGTGTGGAGACACTCGGAGCGGGGCCCCAGGCAAACT 955
DB 1469 CAGTCACCGACAACATGCTGTGTGTGGAGACACTCGGAGCGGGGCCCCAGGCAAACT 1528
QY 956 TGCAGGAGCCTGCCAGGCGATTCGAGAGCGCCCTGCTGTCTGACGATGCGCGCA 1015
DB 1529 TGCAGGAGCCTGCCAGGCGATTCGAGAGCGCCCTGCTGTCTGACGATGCGCGCA 1588
QY 1016 TGACTTTGGTGGGATCATCAGCTGGGGCTGGGCTGTGGACAGAGGATGTCCGGGTG 1075
DB 1589 TGACTTTGGTGGGATCATCAGCTGGGGCTGGGCTGTGGACAGAGGATGTCCGGGTG 1648
QY 1076 TGTCACAAAGGTTTACCAACTACCTAGACTGGATTCGTCGACAAATCGGACCG 1128
DB 1649 TGTCACAAAGGTTTACCAACTACCTAGACTGGATTCGTCGACAAATCGGACCG 1701
```

## RESULT 9

```
US-08-883-795A-39
; Sequence 39, Application US/08883795A
; Patent No. 5985607
```

## GENERAL INFORMATION:

```
; APPLICANT: Delcuve, Genevieve
; APPLICANT: Awang, Gregor
; TITLE OF INVENTION: Recombinant DNA Molecules and Expression
; TITLE OF INVENTION: Vectors for Tissue Plasminogen Activator
; NUMBER OF SEQUENCES: 39
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BERESKIN & PARR
; STREET: 40 King Street West
; CITY: Toronto
; STATE: Ontario
```

```
; COUNTRY: Canada
; ZIP: M5H 3Y2
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/883,795A
; FILING DATE: 27-JUN-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Gravelle, Micheline
; REGISTRATION NUMBER: 40,261
; REFERENCE/DOCKET NUMBER: 7841-062
; TELEPHONE: (416) 364-7311
; TELEPHONE: (416) 361-1398
; INFORMATION FOR SEQ ID NO: 39:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1955 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; US-08-883-795A-39
```

```
Query Match 94.3%; Score 1063.4; DB 2; Length 1955;
Best Local Similarity 99.4%; Pred. No. 2.5e-289;
Matches 1067; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
```

```
QY 56 CCACAGGCGGCTCTGAGGAAACAGTCACTGTCTACTTTGGGAATGGGTACGCTACCGTG 115
DB 704 CCCCTGCTCTCTGAGGAAACAGTCACTGTCTACTTTGGGAATGGGTACGCTACCGTG 763
QY 116 GCACGACAGCCTCACCGAGTCCGGTGCCTCTCCCTCCGTTGGAATTCATGATCCTGA 175
DB 764 GCACGACAGCCTCACCGAGTCCGGTGCCTCTCCCTCCGTTGGAATTCATGATCCTGA 823
QY 176 TAGCAAGGTTTACACAGCACAGAACCCAGTGCACAGGCACTGGGGCTGGGCAAAACATA 235
DB 824 TAGCAAGGTTTACACAGCACAGAACCCAGTGCACAGGCACTGGGGCTGGGCAAAACATA 883
QY 236 ATTACTGCCGAATCTCTGATGGGATGCCAAGCGCTGTCACGCTGTCAGAGAACCGCA 295
DB 884 ATTACTGCCGAATCTCTGATGGGATGCCAAGCGCTGTCACGCTGTCAGAGAACCGCA 943
QY 296 GGCTGACGTGGGAGTACTGTGTCCTCTGCTCCACTCGCGGCTGAGACAGTACA 355
DB 944 GGCTGACGTGGGAGTACTGTGTCCTCTGCTCCACTCGCGGCTGAGACAGTACA 1003
QY 356 GCCAGCCTCAGTTTTCGATCAAAAGAGGAGGCTCTTTCGCCGACATCGCTCCACCCCTGGC 415
DB 1004 GCCAGCCTCAGTTTTCGATCAAAAGAGGAGGCTCTTTCGCCGACATCGCTCCACCCCTGGC 1063
QY 416 AGGCTGCCATCTTTGCCAAGCACAGGAGGTGCGCGGAGAGGGTTCTGTGCGGGGCA 475
DB 1064 AGGCTGCCATCTTTGCCAAGCACAGGAGGTGCGCGGAGAGGGTTCTGTGCGGGGCA 1123
QY 476 TACTCATCAGCTCTGCTGGAATCTCTGCGCGCCACTGCTTCCAGGAGAGGTTTCCGC 535
DB 1124 TACTCATCAGCTCTGCTGGAATCTCTGCGCGCCACTGCTTCCAGGAGAGGTTTCCGC 1183
QY 536 CCCACCACTGACGGTGTCTTGGGAGAAACATACCGGGTGGTCCCTGGGAGGAGGAGC 595
DB 1184 CCCACCACTGACGGTGTCTTGGGAGAAACATACCGGGTGGTCCCTGGGAGGAGGAGC 1243
QY 596 AGAAATTTGAAGTCGAAATATACATTGTCATAGGAATTCGATGACACTTACGACA 655
DB 1244 AGAAATTTGAAGTCGAAATATACATTGTCATAGGAATTCGATGACACTTACGACA 1303
```



```
QY 656 ATGACATTTGCGCTGCTGAGCTGAAATCGGATTCGTCCCGCTGTGCCAGAGAGACGCG 715
Db 1304 ATGACATTTGCGCTGCTGAGCTGAAATCGGATTCGTCCCGCTGTGCCAGAGAGACGCG 1363
QY 716 TGGTCCGACATGTGCTCCCGCGGACCTGCGAGCTGCGGAGTGTG 775
Db 1364 TGGTCCGACATGTGCTCCCGCGGACCTGCGAGCTGCGGAGTGTG 1423
QY 776 AGCTCTCCGCTACCGGCAAGCATGAGGCTTTGTCTCTCTTTCTATTTCGAGCGGCTGAAGG 835
Db 1424 AGCTCTCCGCTACCGGCAAGCATGAGGCTTTGTCTCTCTTTCTATTTCGAGCGGCTGAAGG 1483
QY 836 AGGCTCATGTGAGTGTACCATTCAGCCCTGCGACATCAACAATTTACTTAACAGAA 895
Db 1484 AGGCTCATGTGAGTGTACCATTCAGCCCTGCGACATCAACAATTTACTTAACAGAA 1543
QY 896 CAGTCACCGCAACAATGCTGTGCTGGAGACATCGGAGCGGCGGCGCCAGGCAAACT 955
Db 1544 CAGTCACCGCAACAATGCTGTGCTGGAGACATCGGAGCGGCGGCGCCAGGCAAACT 1603
QY 956 TGCACGACGCTGCGGAGGCGAATTCGGGAGGCGCCCTGCTGTGTCTGAACGATGCGCGCA 1015
Db 1604 TGCACGACGCTGCGGAGGCGAATTCGGGAGGCGCCCTGCTGTGTCTGAACGATGCGCGCA 1663
QY 1016 TGACTTTGGTGGGCAATCATGCTGGGCTTGGGCTGTGGAGAGATGTCGCGGCTG 1075
Db 1664 TGACTTTGGTGGGCAATCATGCTGGGCTTGGGCTGTGGAGAGATGTCGCGGCTG 1723
QY 1076 TGTACACAAAGTTTACCAACTTACCTAGACTGGATTCGTGACAACTGCGACG 1128
Db 1724 TGTACACAAAGTTTACCAACTTACCTAGACTGGATTCGTGACAACTGCGACG 1776

RESULT 10
5344773-1
; Patent No. 5344773
; APPLICANT: WEI, CHA-MER; HSIUNG, NANCY; REDDY, VERMURI B.;
; LEMONT, JEFFREY F.; DACKOWSKI, WILLIAM; DOUGLAS, RICHARD;
; COLE, EDWARD S.; PURCELL JR., RICHARD D.; LAU, DAVID TAI-YUI
; TITLE OF INVENTION: HUMAN UTERINE TISSUE PLASMINOGEN
; ACTIVATOR PRODUCED BY RECOMBIANT DNA
; NUMBER OF SEQUENCES: 6
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/782,686
; FILING DATE: 01-OCT-1985
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 656,770
; FILING DATE: 01-OCT-1984
; SEQ ID NO:1:
; LENGTH: 2457
5344773-1

Query Match 94.3%; Score 1063.4; DB 6; Length 2457;
Best Local Similarity 99.4%; Pred. No. 2.7e-289;
Matches 1067; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 56 CCCAGGCGGCTCTGAGGAAACAGTACTCTACTTTGGGAATGGGTTCAGCTTACCGTG 115
Db 627 CCCCTGCTGCTGAGGAAACAGTACTCTACTTTGGGAATGGGTTCAGCTTACCGTG 686
QY 116 GCACGACAGCTTACAGCAGACAGAACCCAGTCCGAGGACCTGGGCTGGGCAACATA 175
Db 687 GCACGACAGCTTACAGCAGACAGAACCCAGTCCGAGGACCTGGGCTGGGCAACATA 746
QY 176 TAGGCAAGTTTACAGCAGACAGAACCCAGTCCGAGGACCTGGGCTGGGCAACATA 235
Db 747 TAGGCAAGTTTACAGCAGACAGAACCCAGTCCGAGGACCTGGGCTGGGCAACATA 806
QY 236 ATTACTGCGGAATCTTGTATGGGATGCCAAGCCCTGTGTGACGTGCTGAAGAACGCA 295
Db 807 ATTACTGCGGAATCTTGTATGGGATGCCAAGCCCTGTGTGACGTGCTGAAGAACGCA 866
QY 296 GGCTGACGTGGGAGTACTGTGATGCGCTCTCTGCTCCACCTGCGGCTGAGACAGTACA 355
```

```
Db 867 GGCTGACGTGGGAGTACTGTGATGTGCCCTCTCTGCTCCACCTGCGGCTGAGACAGTACA 926
QY 356 GCACGCTCAGTTTTCGATCAAAAGGAGGCTCTTTCCGCGGACATCGCTCCACCCCTGGC 415
Db 927 GCACGCTCAGTTTTCGATCAAAAGGAGGCTCTTTCCGCGGACATCGCTCCACCCCTGGC 986
QY 416 AGGCTGCATCTTTTGCACAGCAGAGGCTGCGCCGAGAGCGGTTCCTGTGCGGGGCA 475
Db 987 AGGCTGCATCTTTTGCACAGCAGAGGCTGCGCCGAGAGCGGTTCCTGTGCGGGGCA 1046
QY 476 TACTCATCAGTCTGCTGATTTCTCTGCGGCGGCTCTGCTCCAGAGAGGTTTCGCG 535
Db 1047 TACTCATCAGTCTGCTGATTTCTCTGCGGCGGCTCTGCTCCAGAGAGGTTTCGCG 1106
QY 536 CCACACACCTGAGCGTGATCTTTGGGAGAAACATACCGGCTGGTCCCTGGGAGAGAGC 595
Db 1107 CCACACACCTGAGCGTGATCTTTGGGAGAAACATACCGGCTGGTCCCTGGGAGAGAGC 1166
QY 596 AGAAATTTGAAGTCGAAAAATACATTTGTCATAAGGAATTCGATGATGACATTTACGACA 655
Db 1167 AGAAATTTGAAGTCGAAAAATACATTTGTCATAAGGAATTCGATGATGACATTTACGACA 1226
QY 656 ATGACATTTGCGCTGCTGAGCTGAAATCGGATTCGCTCCGCTGTGCCAGAGAGCAGCG 715
Db 1227 ATGACATTTGCGCTGCTGAGCTGAAATCGGATTCGCTCCGCTGTGCCAGAGAGCAGCG 1286
QY 716 TGGTCCGACATGTGCTCTTCCCGCGGACCTGCGAGCTGCGGAGCTGGAACGAGTGTG 775
Db 1287 TGGTCCGACATGTGCTCTTCCCGCGGACCTGCGAGCTGCGGAGTGTG 1346
QY 776 AGCTCTCCGCTACGCGCAAGCATGAGGCTTTGCTCTCTTCTATTTCGAGCGGCTGAAGG 835
Db 1347 AGCTCTCCGCTACGCGCAAGCATGAGGCTTTGCTCTCTTCTATTTCGAGCGGCTGAAGG 1406
QY 836 AGGCTCATGTGAGCTGTATCCATCCAGCGCTGCAATCAACAATTTACTTAACAGAA 895
Db 1407 AGGCTCATGTGAGCTGTATCCATCCAGCGCTGCAATCAACAATTTACTTAACAGAA 1466
QY 896 CAGTCACCGCAACATGCTGTGCTGGAGACATCGGAGCGGCGGCGCCAGGCAAACT 955
Db 1467 CAGTCACCGCAACATGCTGTGCTGGAGACATCGGAGCGGCGGCGCCAGGCAAACT 1526
QY 956 TGCACGACGCTGCGGAGGCGAATTCGGGAGGCGCCCTGCTGTGCTGAACGATGCGCGCA 1015
Db 1527 TGCACGACGCTGCGGAGGCGAATTCGGGAGGCGCCCTGCTGTGCTGAACGATGCGCGCA 1586
QY 1016 TGACTTTGGTGGGCAATCATCAGCTGGGCTGTGGGCTGTGGAGAGGATGTCGCGGCTG 1075
Db 1587 TGACTTTGGTGGGCAATCATCAGCTGGGCTGTGGGCTGTGGAGAGGATGTCGCGGCTG 1646
QY 1076 TGTACACAAAGTTTACCAACTTACCTAGACTGGATTCGTGACAACTGCGGCGG 1128
Db 1647 TGTACACAAAGTTTACCAACTTACCTAGACTGGATTCGTGACAACTGCGGCGG 1699

RESULT 11
US-08-740-1
; Sequence 1, Application US/08286740
; Patent No. 5561053
; GENERAL INFORMATION:
; APPLICANT: Crowley, Craig W.
; TITLE OF INVENTION: METHOD FOR SELECTING HIGH-EXPRESSING
; TITLE OF INVENTION: HOST CELLS
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESS: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
```

MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: patin (Genentech)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/286.740

FILING DATE: 05-AUG-1994

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Lee, Wendy M.

REGISTRATION NUMBER:

REFERENCE/DOCKET NUMBER: 798

TELEPHONE: 415/225-1994

TELEFAX: 415/952-9881

TELEX: 910/371-7168

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 7360 bases

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

US-08-286-740-1

Query Match 94.3%; Score 1063.4; DB 1; Length 7360;

Best Local Similarity 99.4%; Pred. No. 4.3e-289;

Matches 1067; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```
QY 56 CCCAGCGCGCTCTCAGGGAACAGTACTTCTTTGGGAATGGGTTCAGCTACCGTG 115
DB 2235 CCCCTGCTCTCTCAGGGAACAGTACTTCTTTGGGAATGGGTTCAGCTACCGTG 2294
QY 116 GCACGCACAGCTCTCAGGAGTGGGTCTCTCTCCCTCCGCTGGAAATTCATGATCTGA 175
DB 2295 GCACGCACAGCTCTCAGGAGTGGGTCTCTCTCCCTCCGCTGGAAATTCATGATCTGA 2354
QY 176 TAGGCAAGTTTACACAGACAGAACCCAGTCCAGGCACTGGGCTGGGCAAAACATA 235
DB 2355 TAGGCAAGTTTACACAGACAGAACCCAGTCCAGGCACTGGGCTGGGCAAAACATA 2414
QY 236 ATTACTCGCGGAATCTCATGGGATGCCAGCCCTGTGTGCACGTCTGAAGAACGCA 295
DB 2415 ATTACTCGCGGAATCTCATGGGATGCCAGCCCTGTGTGCACGTCTGAAGAACGCA 2474
QY 296 GGCTGACGTGGGAGTACTGTGATGTGCCCTCTCTCTCCACCTGGCGCTGAGACAGTACA 355
DB 2475 GGCTGACGTGGGAGTACTGTGATGTGCCCTCTCTCTCCACCTGGCGCTGAGACAGTACA 2534
QY 356 GCCAGCTCAGTTTCGATCAAGAGAGGCTCTTCGCGCAATCGCCTCCACCCCTGGC 415
DB 2535 GCCAGCTCAGTTTCGATCAAGAGAGGCTCTTCGCGCAATCGCCTCCACCCCTGGC 2594
QY 416 AGGCTGCCATCTTTGCCAAGACAGAGAGTCCGCGGAGAGCGTTCCTGCGGGGCA 475
DB 2595 AGGCTGCCATCTTTGCCAAGACAGAGAGTCCGCGGAGAGCGTTCCTGCGGGGCA 2654
QY 476 TACTCATCAGCTCTCTGCTGGATTTCTCTCTGCGGCCACTGTCTTCAGGAGAGTTTCGCG 535
DB 2655 TACTCATCAGCTCTCTGCTGGATTTCTCTCTGCGGCCACTGTCTTCAGGAGAGTTTCGCG 2714
QY 536 CCCACCACCTGACGCTGATCTTTGGCGCAATACCGGGTGGTCTCTGGCGAGAGGAGC 595
DB 2715 CCCACCACCTGACGCTGATCTTTGGCGCAATACCGGGTGGTCTCTGGCGAGGAGGAGC 2774
QY 596 AGAATTTGAATTCGAAATATACATTGTCCTAAGGAATTCGATGATCAGACTTACGACA 655
DB 2775 AGAATTTGAATTCGAAATATACATTGTCCTAAGGAATTCGATGATCAGACTTACGACA 2834
QY 656 ATGACATTGCGTCTGTCAGCTGAAATCGGATTCGTCCTCGCTGTGCCAGGAGAGCAGC 715
```

```
DB 2835 ATGACATTGCGTCTGTCAGCTGAAATCGGATTCGTCCTCCGCTGTCGCCAGAGACGCG 2894
QY 716 TGGTCCGCACTGTGTGCTTCCCGCGGAGACCTGCGAGCTGCCGAGCTGGACGAGTGTG 775
DB 2895 TGGTCCGCACTGTGTGCTTCCCGCGGAGACCTGCGAGCTGCCGAGCTGGACGAGTGTG 2954
QY 776 AGCTCTCCGCTACGGCAAGCATGAGGCTTGTCTCTCTTTCTATTTCGAGCGGCTGAAGG 835
DB 2955 AGCTCTCCGCTACGGCAAGCATGAGGCTTGTCTCTCTTTCTATTTCGAGCGGCTGAAGG 3014
QY 836 AGGCTCATGTGACAGTGTATCCCATCCAGCGCTGCACATCACAACTTTACTTAACAGAA 895
DB 3015 AGGCTCATGTGACAGTGTATCCCATCCAGCGCTGCACATCACAACTTTACTTAACAGAA 3074
QY 896 CAGTCACCGCAACATGCTGTGTCTGGAGACACTCGGAGCGGCGGCCCCAGGCAACT 955
DB 3075 CAGTCACCGCAACATGCTGTGTCTGGAGACACTCGGAGCGGCGGCCCCAGGCAACT 3134
QY 956 TGACACGACGCTGCGAGGCGGATTCGGGAGGCGCCCTGGTGTGTCTGAACGATGCCGCA 1015
DB 3135 TGACACGACGCTGCGAGGCGGATTCGGGAGGCGCCCTGGTGTGTCTGAACGATGCCGCA 3194
QY 1016 TGACTTTGGTGGGCACTCATCAGCTGGGCGCTGGGCTGTGGACAGAGATGTCGCGGTG 1075
DB 3195 TGACTTTGGTGGGCACTCATCAGCTGGGCGCTGGGCTGTGGACAGAGATGTCGCGGTG 3254
QY 1076 TGATACAAAGTTTACCACTACCTAGACTGGATTCGTCGACAACTGCGACG 1128
DB 3255 TGATACAAAGTTTACCACTACCTAGACTGGATTCGTCGACAACTGCGACG 3307
```

# RESULT 12

PCT-US95-09576-1

; Sequence 1, Application PC/TUS9509576

; GENERAL INFORMATION:

; APPLICANT: GENENTECH, INC.

; TITLE OF INVENTION: METHOD FOR SELECTING HIGH-EXPRESSING

; TITLE OF INVENTION: HOST CELLS

; NUMBER OF SEQUENCES: 4

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Genentech, Inc.

; STREET: 460 Point San Bruno Blvd

; CITY: South San Francisco

; STATE: California

; COUNTRY: USA

; ZIP: 94080

; COMPUTER READABLE FORM:

; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: patin (Genentech)

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: PCT/US95/09576

; FILING DATE:

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/286740

; FILING DATE: 05-AUG-1994

; ATTORNEY/AGENT INFORMATION:

; NAME: Lee, Wendy M.

; REFERENCE/DOCKET NUMBER: 798PCT

; TELEPHONE: 415/225-1994

; TELEFAX: 415/952-9881

; TELEX: 910/371-7168

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 7360 bases

; TYPE: nucleic acid

; STRANDEDNESS: double

; TOPOLOGY: linear

PCT-US95-09576-1

Query Match 94.3%; Score 1063.4; DB 5; Length 7360;  
Best Local Similarity 99.4%; Pred. No. 4.3e-289;  
Matches 1067; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 56 CCCAGCGGCTCTCAGGGAACAGTGTCTCTTTGGGAATGGTTCAGCTACCGTG 115  
DB 2235 CCCCTGCTCTCTCAGGGAACAGTGTCTCTTTGGGAATGGTTCAGCTACCGTG 2294

QY 116 GCAGCAGAGCTCACCAGTGGGTGCTCTCTGCTCCCGTGAATTCATGATCTGA 175  
DB 2295 GCAGCAGAGCTCACCAGTGGGTGCTCTCTGCTCCCGTGAATTCATGATCTGA 2354

QY 176 TAGGCAAGTTTACACAGACAGAACCCAGTGGCCAGTGGCTGGGCAAAACATA 235  
DB 2355 TAGGCAAGTTTACACAGACAGAACCCAGTGGCCAGTGGCTGGGCAAAACATA 2414

QY 236 ATTACTGCGGGAATCCTGATGGGATCCAGAGCCCTGGTGCACGTGTGAAGACCGCA 295  
DB 2415 ATTACTGCGGGAATCCTGATGGGATCCAGAGCCCTGGTGCACGTGTGAAGACCGCA 2474

QY 296 GGCTGACGTGGAGTACTGTGATGTGCCCTCTCTGCTCCACCTGGCGCTGAGACAGTACA 355  
DB 2475 GGCTGACGTGGAGTACTGTGATGTGCCCTCTCTGCTCCACCTGGCGCTGAGACAGTACA 2534

QY 356 GCCAGCCTCAGTTTCGATCAAGAGGGCTCTTCCGCCGACATCGCTCCCAACCCCTGGC 415  
DB 2535 GCCAGCCTCAGTTTCGATCAAGAGGGCTCTTCCGCCGACATCGCTCCCAACCCCTGGC 2594

QY 416 AGGCTGCCATCTTTGCCAGACAGAGGTGGCCGGAGGGTTCCTGTGGGGGCA 475  
DB 2595 AGGCTGCCATCTTTGCCAGACAGAGGTGGCCGGAGGGTTCCTGTGGGGGCA 2654

QY 476 TACTCATCAGTCTCTGCTGGATTTCTCTGCGGCCACTGCTTCAGGAGAGTTTCGC 535  
DB 2655 TACTCATCAGTCTCTGCTGGATTTCTCTGCGGCCACTGCTTCAGGAGAGTTTCGC 2714

QY 536 CCCACCCTGACGGTGTATTTGGGCAAAACATACCGGGTGGTCTCTGGCAGGAGGAGC 595  
DB 2715 CCCACCCTGACGGTGTATTTGGGCAAAACATACCGGGTGGTCTCTGGCAGGAGGAGC 2774

QY 596 AGAATTTGAAGTGGAAATACATTTGCCATAGGAATTCGATGATGACATTTACGACA 655  
DB 2775 AGAATTTGAAGTGGAAATACATTTGCCATAGGAATTCGATGATGACATTTACGACA 2834

QY 656 ATGACATTTGCTGCTGAGCTGAAATCGGATTCGTCCCGCTGTGCCAGGAGACGCG 715  
DB 2835 ATGACATTTGCTGCTGAGCTGAAATCGGATTCGTCCCGCTGTGCCAGGAGACGCG 2894

QY 716 TGGTCCGCACTGTGTGCTTCCCGCGGACCTGCACTGCGGACTGGACGGAGTGTG 775  
DB 2895 TGGTCCGCACTGTGTGCTTCCCGCGGACCTGCACTGCGGACTGGACGGAGTGTG 2954

QY 776 AGCTTCGGCTACGGCAAGCATGAGCGCTTGTCTCTTTCTATTCGGAGCGGTGAAG 835  
DB 2955 AGCTTCGGCTACGGCAAGCATGAGCGCTTGTCTCTTTCTATTCGGAGCGGTGAAG 3014

QY 836 AGGCTCATGTGACACTGTACCATCCAGCGCTGCACATCAACATTTACTTAACAGAA 895  
DB 3015 AGGCTCATGTGACACTGTACCATCCAGCGCTGCACATCAACATTTACTTAACAGAA 3074

QY 896 CAGTCAACCGAACATGCTGTGTGTGGAGACCTCGGAGCGGGGGCCCGAGCAAACT 955  
DB 3075 CAGTCAACCGAACATGCTGTGTGTGGAGACCTCGGAGCGGGGGCCCGAGCAAACT 3134

QY 956 TGCAGGACCTGCCAGGCGCATTCGGAGAGCCCTGTGTGTCTGAACGATGGCCGCA 1015  
DB 3135 TGCAGGACCTGCCAGGCGCATTCGGAGAGCCCTGTGTGTCTGAACGATGGCCGCA 3194

QY 1016 TGACTTTGGTGGGATCATCAGCTGGGCGCTGGGCTGTGGACAGAGGATGTCCCGGTG 1075  
DB 3195 TGACTTTGGTGGGATCATCAGCTGGGCGCTGGGCTGTGGACAGAGGATGTCCCGGTG 3254

QY 1076 TGTACACAAGGTTACCAACTACCTAGACTGGATTGCTGACAAACATGCGACCG 1128  
DB 3255 TGTACACAAGGTTACCAACTACCTAGACTGGATTGCTGACAAACATGCGACCG 3307

RESULT 13  
US-08-811-949-44  
; Sequence 44, Application US/08811949  
; Patent No. 5840533  
; GENERAL INFORMATION:  
; APPLICANT: NIWA, MINEO  
; APPLICANT: SAITO, YOSHIMASA  
; APPLICANT: SASAKI, HITOSHI  
; APPLICANT: HAYASHI, MASAKO  
; APPLICANT: NOTANI, JOUJI  
; APPLICANT: KOBAYASHI, MASAKAZU  
; TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR  
; NUMBER OF SEQUENCES: 67  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,  
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400  
; CITY: ARLINGTON  
; STATE: VA  
; COUNTRY: USA  
; ZIP: 22202  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/811,949  
; FILING DATE: 05-MAR-1997  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: OBLON, NORMAN F.  
; REGISTRATION NUMBER: 24,618  
; REFERENCE/DOCKET NUMBER: 18-966-0  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 703-413-3000  
; TELEFAX: 703-413-2220  
; INFORMATION FOR SEQ ID NO: 44:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1068 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: circular  
; MOLECULE TYPE: DNA (genomic)  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 1..1065  
US-08-811-949-44

Query Match 94.1%; Score 1062; DB 2; Length 1068;  
Best Local Similarity 100.0%; Pred. No. 4.8e-289;  
Matches 1062; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 67 TCTGAGGGAACAGTGTCTCTGCTCCCGTGAATTCATGATTCCTGATAGGCAAGTT 186  
DB 4 TCTGAGGGAACAGTGTCTCTGCTCCCGTGAATTCATGATTCCTGATAGGCAAGTT 63

QY 127 CTCAACAGTGGGTGCTCTGCTCCCGTGAATTCATGATTCCTGATAGGCAAGTT 186  
DB 64 CTCAACAGTGGGTGCTCTGCTCCCGTGAATTCATGATTCCTGATAGGCAAGTT 123

QY 187 TACACAGCAGAACCCAGTGGCCAGGCACTGGGCTGGGCAAAACATAATTAATGCGCG 246  
DB 124 TACACAGCAGAACCCAGTGGCCAGGCACTGGGCTGGGCAAAACATAATTAATGCGCG 183

QY 247 AATCTGATGGGATGCCAAGCCCTGTGTGCCAAGTGTGAGAACCCAGCTGACGTGG 306  
DB 184 AATCTGATGGGATGCCAAGCCCTGTGTGCCAAGTGTGAGAACCCAGCTGACGTGG 243

QY 307 GAGTACTGTGATGTCCTCTGCTCCACCTGGCCCTGAGACATGACGACGCTCAG 366  
DB 244 GAGTACTGTGATGTCCTCTGCTCCACCTGGCCCTGAGACATGACGACGCTCAG 303  
QY 367 TTTGTCATCAAGAGGAGCTTTGCGCAGATCGCTCCACCCCTGGCAGGCTGCCATC 426  
DB 304 TTTGTCATCAAGAGGAGCTTTGCGCAGATCGCTCCACCCCTGGCAGGCTGCCATC 363  
QY 427 TTTGTCATCAAGAGGAGCTTTGCGCAGATCGCTCCACCCCTGGCAGGCTGCCATC 486  
DB 364 TTTGTCATCAAGAGGAGCTTTGCGCAGATCGCTCCACCCCTGGCAGGCTGCCATC 423  
QY 487 TCCTGCTGATCTCTCTGCGCCCACTGCTTCCAGGAGAGGTTTCCGCCCCCAACACCTG 546  
DB 424 TCCTGCTGATCTCTCTGCGCCCACTGCTTCCAGGAGAGGTTTCCGCCCCCAACACCTG 483  
QY 547 ACCTGTGATCTTTGGCAGAACATACCGGTGTGCTTCCAGGAGGAGGACAGAAATTTGAA 606  
DB 484 ACCTGTGATCTTTGGCAGAACATACCGGTGTGCTTCCAGGAGGAGGACAGAAATTTGAA 543  
QY 607 GTCGAAATACATTTGTCATTAAGAAATTCGATGATGACACTTACGCAATGACATTCG 666  
DB 544 GTCGAAATACATTTGTCATTAAGAAATTCGATGATGACACTTACGCAATGACATTCG 603  
QY 667 CTGCTGACGCTGAATCGGATTCGCTCCGCTGTGCCAGGAGAGCAGCGTGTGCGCACT 726  
DB 604 CTGCTGACGCTGAATCGGATTCGCTCCGCTGTGCCAGGAGAGCAGCGTGTGCGCACT 663  
QY 727 GTGTGCTTCCCGCGGAGACTGCGAGTCCGAGCTGCGGAGCTGAGCGAGTGTGAGCTTCCGCG 786  
DB 664 GTGTGCTTCCCGCGGAGACTGCGAGTCCGAGCTGCGGAGCTGAGCGAGTGTGAGCTTCCGCG 723  
QY 787 TACGCAAGCATGAGGCTTGTCTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 846  
DB 724 TACGCAAGCATGAGGCTTGTCTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 783  
QY 847 AGACTGTACCATCCAGCGCTGACATCAACATTTACTTTAAACAGAAAGTCAACGAC 906  
DB 784 AGACTGTACCATCCAGCGCTGACATCAACATTTACTTTAAACAGAAAGTCAACGAC 843  
QY 907 AACATGTGTGTGTTGAGACACTCGGAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 966  
DB 844 AACATGTGTGTGTTGAGACACTCGGAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 903  
QY 967 TCCAGGCGGATTCGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 1026  
DB 904 TCCAGGCGGATTCGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 963  
QY 1027 GGCATCATCAGCTGGGCGCTGGGCTGTGGACAGAGGATGTCCCGGGTGTGTACACAAAG 1086  
DB 964 GGCATCATCAGCTGGGCGCTGGGCTGTGGACAGAGGATGTCCCGGGTGTGTACACAAAG 1023  
QY 1087 GTTACCAACTACCTAGACTGATGATGTCGACAAACATGCGACCG 1128  
DB 1024 GTTACCAACTACCTAGACTGATGATGTCGACAAACATGCGACCG 1065

## RESULT 14

US-08-811-949-50  
; Sequence 50, Application US/08811949  
; Patent No. 5840533  
; GENERAL INFORMATION:  
; APPLICANT: NIWA, MINEO  
; APPLICANT: SAITO, YOSHIMASA  
; APPLICANT: SASAKI, HITOSHI  
; APPLICANT: HAYASHI, MASAKO  
; APPLICANT: NOTANI, JOUJI  
; APPLICANT: KOBAYASHI, MASAKAZU  
; TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR  
; NUMBER OF SEQUENCES: 67  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,

; ADDRESSEE: P.C.  
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400  
; CITY: ARLINGTON  
; STATE: VA  
; COUNTRY: USA  
; ZIP: 22202  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/811,949  
; FILING DATE: 05-MAR-1997  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: OBLON, NORMAN F.  
; REGISTRATION NUMBER: 24,618  
; REFERENCE/DOCKET NUMBER: 18-966-0  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 703-413-3000  
; TELEFAX: 703-413-2220  
; INFORMATION FOR SEQ ID NO: 50:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1314 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: circular  
; MOLECULE TYPE: DNA (genomic)  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 1..1311  
; US-08-811-949-50

Query Match 94.1%; Score 1061.8; DB 2; Length 1314;

Best Local Similarity 99.3%; Pred. No. 6e-289;  
Matches 1066; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 56 CCCAGCGGCTCTGAGGAAACAGTCACTCTTGGGAAATGGGTCAACCTACCGTG 115  
DB 239 CCCCTGCTCTCTGAGGAAACAGTCACTCTTGGGAAATGGGTCAACCTACCGTG 298  
QY 116 GCACGACAGCCTCACCGAGTCCGCT 175  
DB 299 GCACGACAGCCTCACCGAGTCCGCT 358  
QY 176 TAGCAAGGTTTACAGACACAGAAACCCAGTCCAGGCACTGGGCTGGGCAACATA 235  
DB 359 TAGCAAGGTTTACAGACACAGAAACCCAGTCCAGGCACTGGGCTGGGCAACATA 418  
QY 236 ATTACTGCGGAATCTGATGGGATGCCAAGCCCTGGTCCACGCTGAGAACCGCA 295  
DB 419 ATTACTGCGGAATCTGATGGGATGCCAAGCCCTGGTCCACGCTGAGAACCGCA 478  
QY 296 GGCTGACGTGGGAGTACTGTGATGTGCT 355  
DB 479 GGCTGACGTGGGAGTACTGTGATGTGCT 538  
QY 356 GCACGCTCAGTTTTCGATCAAAAGGAGGCTCTTTCGCGGACATGCGCTCCACCCCTGGC 415  
DB 539 GCCAGCCTCAGTTTTCGATCAAAAGGAGGCTCTTTCGCGGACATGCGCTCCACCCCTGGC 598  
QY 416 AGGCTGCTCATCTTTGCCAAGCAGAGGAGTCCGCCGAGAGCGGTCTCTGTGCGGGGCA 475  
DB 599 AGGCTGCTCATCTTTGCCAAGCAGAGGAGTCCGCCGAGAGCGGTCTCTGTGCGGGGCA 658  
QY 476 TACTCATCAGCTCCTGTGATTTCTCTCTGCGGCCCACTGCTTCCAGGAGGTTTCCGC 535  
DB 659 TACTCATCAGCTCCTGTGATTTCTCTCTGCGGCCCACTGCTTCCAGGAGGTTTCCGC 718  
QY 536 CCACACACCTGACGCTGATCTTGGGCAAGACATACCGGGTGGTCCCTGGGAGAGGAGC 595  
DB 719 CCACACACCTGACGCTGATCTTGGGCAAGACATACCGGGTGGTCCCTGGGAGGAGGAGC 778

596	QY	AGAAATTTGAAGTCGAAAAATACATTGTCCTCATAGGAATTCGATGATGACACTTACGACA	655
779	Db	AGAAATTTGAAGTCGAAAAATACATTGTCCTCATAGGAATTCGATGATGACACTTACGACA	838
656	QY	ATGACATTGCGCTGCTGCAGCTGAAATTCGGATTGTCCTCCGCTGTGCCGCCAGGAGACGCG	715
839	Db	ATGACATTGCGCTGCTGCAGCTGAAATTCGGATTGTCCTCCGCTGTGCCGCCAGGAGACGCG	898
716	QY	TGCTCCGCACTGTGTGCTTCCCGCGGCACTGCAGCTGCCGGACTGACGGAGTGTG	775
899	Db	TGCTCCGCACTGTGTGCTTCCCGCGGCACTGCAGCTGCCGGACTGACGGAGTGTG	958
776	QY	AGCTCTCCGGCTACGGCAAGCATGAGGCTTTGCTCCTTTCTATTCCGAGCGGCTGAAGG	835
959	Db	AGCTCTCCGGCTACGGCAAGCATGAGGCTTTGCTCCTTTCTATTCCGAGCGGCTGAAGG	1018
836	QY	AGSCTCATGTCAGACTGTATCCCATCCAGCGCTGCAATCACAACTTTACTTTAAACAGAA	895
1019	Db	AGSCTCATGTCAGACTGTATCCCATCCAGCGCTGCAATCACAACTTTACTTTAAACAGAA	1078
896	QY	CAGTCAACCGACACATGCTGTGTCTGGAGACACTCGGAGCGGGGCCCCAGCGCAAACT	955
1079	Db	CAGTCAACCGACACATGCTGTGTCTGGAGACACTCGGAGCGGGGCCCCAGCGCAAACT	1138
956	QY	TGCACAGACGCTGCACAGGGCGATTCCGGAGGCCCCCTGGTGCTCTGAACGATGGCCGCA	1015
1139	Db	TGCACAGACGCTGCACAGGGCGATTCCGGAGGCCCCCTGGTGCTCTGAACGATGGCCGCA	1198
1016	QY	TGACTTTTGGTGGGCATCATCAGCTGGGGCCTGGGCTGTGGACAGAGGATGTGCCGGGTG	1075
1199	Db	TGACTTTTGGGSCATCATCAGCTGGGGCCTGGGCTGTGGACAGAGGATGTGCCGGGTG	1258
1076	QY	TGTACAAAGGTTACCACTACTAGCTGGATTCTGTGACAAATCGCGACCG	1128
1259	Db	TGTACAAAGGTTACCACTACTAGCTGGATTCTGTGACAAATCGCGACCG	1311

RESULT 15

US-08-811-949-38  
Sequence 38, Application US/08811949  
Patent No. 5840533  
GENERAL INFORMATION:  
APPLICANT: NIWA, MINEO  
APPLICANT: SAITO, YOSHIMASA  
APPLICANT: SASAKI, HITOSHI  
APPLICANT: HAYASHI, MASAKO  
APPLICANT: NOTANI, JOUJI  
APPLICANT: KOBAYASHI, MASAKAZU  
TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR  
NUMBER OF SEQUENCES: 67  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
ADDRESSES: P.C.  
STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400  
CITY: ARLINGTON  
STATE: VA  
COUNTRY: USA  
ZIP: 22202  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC Compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/811,949  
FILING DATE: 05-MAR-1997  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: OBLON, NORMAN F.  
REGISTRATION NUMBER: 24,618  
REFERENCE/DOCKET NUMBER: 18-966-0  
TELECOMMUNICATION INFORMATION:

Qy 896 CAGTCACCGACAACATGCTGTGTGCTGGAGACACTCGGAGCGGGGCCCCAGGCAAACT 955  
|||  
Db 1351 CAGTCACCGACAACATGCTGTGTGCTGGAGACACTCGGAGCGGGGCCCCAGGCAAACT 1410  
|||  
Qy 956 TGCACGACGCCCTGCCAGGGCGATTTCGGGAGGGCCCCCTGGTGTGTCTGAAACGATGGCCGCA 1015  
|||  
Db 1411 TGCACGACGCCCTGCCAGGGCGATTTCGGGAGGGCCCCCTGGTGTGTCTGAAACGATGGCCGCA 1470  
|||  
Qy 1016 TGACTTTGGTGGGCATCATCAGCTGGGGCCTGGGCTGTGGACAGAGGATGTCGGGGTG 1075  
|||  
Db 1471 TGACTTTGGTGGGCATCATCAGCTGGGGCCTGGGCTGTGGACAGAGGATGTCGGGGTG 1530  
|||  
Qy 1076 TGTACACAAGGTTACCAACTACCTAGACTGGATTTCGTGACAACATGGGACCG 1128  
|||  
Db 1531 TGTACACAAGGTTACCAACTACCTAGACTGGATTTCGTGACAACATGGGACCG 1583  
|||

Search completed: August 19, 2003, 08:35:10  
Job time : 89.2804 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 19, 2003, 04:46:44 ; Search time 318.307 Seconds  
(without alignments)  
7930.701 Million cell updates/sec

Title: US-09-987-455-5  
Perfect score: 1128  
Sequence: 1 atgaaaagacagctatgc.....tttgtgacaacatgcgacg 1128

Scoring table: IDENTITY\_NUC  
Gap 10.0 , Gapext 1.0

Searched: 1504479 seqs, 1118970152 residues  
Total number of hits satisfying chosen parameters: 3008958

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications NA:\*

- 1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq:\*
- 3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:\*
- 4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:\*
- 5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:\*
- 6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq:\*
- 7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:\*
- 8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*
- 9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq:\*
- 10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq:\*
- 11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq:\*
- 12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:\*
- 13: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq:\*
- 14: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq:\*
- 15: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:\*
- 16: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:\*
- 17: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1128	100.0	1128	11	US-09-987-455-2
2	1128	100.0	1128	11	US-09-987-455-5
3	1063.4	94.3	1689	9	US-09-969-271-6
4	1063.4	94.3	2509	14	US-10-193-656-7
5	1063.4	94.3	2519	9	US-09-969-271-5
6	1062	94.1	1065	11	US-09-987-455-4
7	1062	94.1	1065	11	US-09-987-455-7
8	1061.8	94.1	2641	10	US-09-974-298-144
9	227.2	20.1	329	12	US-10-007-926A-433
10	191.6	17.0	1212	10	US-09-880-503-15
11	188	16.7	1415	14	US-10-198-846-12748
12	170.8	15.1	1475	9	US-09-735-705-122
13	170.8	15.1	1475	10	US-09-850-716A-122
14	170.8	15.1	1475	10	US-09-897-778-122
15	170.8	15.1	1475	11	US-09-466-396A-122
16	170.8	15.1	1475	12	US-10-117-982-122

17	170.8	15.1	1475	12	US-10-101-510-159	Sequence 159, App
18	170.8	15.1	2336	12	US-09-971-392-46	Sequence 46, Appl
19	170.8	15.1	2336	12	US-10-101-510-458	Sequence 458, App
20	167.6	14.9	1296	13	US-10-076-421-1	Sequence 1, Appli
21	167.6	14.9	2294	9	US-09-735-705-123	Sequence 123, App
22	167.6	14.9	2294	10	US-09-850-716A-123	Sequence 123, App
23	167.6	14.9	2294	10	US-09-897-778-123	Sequence 123, App
24	167.6	14.9	2294	11	US-09-466-396A-123	Sequence 123, App
25	167.6	14.9	2294	12	US-10-117-982-123	Sequence 123, App
26	167.6	14.9	2344	12	US-10-101-510-383	Sequence 383, App
27	166	14.7	1236	10	US-09-880-503-12	Sequence 12, Appl
28	166	14.7	2304	12	US-10-301-822-160	Sequence 160, Appl
29	166	14.7	2304	14	US-10-171-311-183	Sequence 183, App
30	163	14.5	474	9	US-09-864-761-1682	Sequence 1989, Ap
31	162	14.4	2486	14	US-10-106-698-1989	Sequence 14, Appl
32	159.8	14.2	831	10	US-09-880-503-14	Sequence 16, Appl
33	159.8	14.2	972	10	US-09-880-503-16	Sequence 16, Appl
34	139	12.3	461	11	US-09-918-995-10405	Sequence 10405, A
35	123	10.9	1959	12	US-10-101-510-743	Sequence 743, App
36	123	10.9	2035	12	US-10-101-510-448	Sequence 1, Appli
37	123	10.9	2048	11	US-09-858-909-1	Sequence 552, App
38	120.4	10.7	2036	10	US-09-954-456-552	Sequence 1612, Ap
39	120.4	10.7	2036	10	US-09-880-107-1612	Sequence 18438, A
40	111.4	9.9	116	9	US-09-864-761-18438	Sequence 11188, A
41	105	9.3	457	11	US-09-918-995-11188	Sequence 237, App
42	89.4	7.9	372	10	US-09-920-300A-237	Sequence 237, App
43	88.4	7.9	372	13	US-10-033-528-237	Sequence 12715, A
44	88	7.8	442	11	US-09-918-995-12715	Sequence 35, Appl
45	83.8	7.4	1230	9	US-09-879-792-35	

ALIGNMENTS

RESULT 1

US-09-987-455-2  
; Sequence 2, Application US/09987455  
; Publication No. US20030049729A1  
; GENERAL INFORMATION:  
; APPLICANT: Aranya Manosroi  
; APPLICANT: Jiradej Manosroi  
; APPLICANT: Chatchai Tayapiwatana  
; APPLICANT: Friedl Guenther Goetz  
; APPLICANT: Rolf-Guenther Werner  
; TITLE OF INVENTION: Methods for Large Scale Production of Recombinant  
; TITLE OF INVENTION: DNA-Derived tPA or K2S Molecules  
; FILE REFERENCE: 0652.2190001  
; CURRENT APPLICATION NUMBER: US/09/987,455  
; CURRENT FILING DATE: 2001-11-14  
; PRIOR APPLICATION NUMBER: 60/268,574  
; PRIOR FILING DATE: 2001-02-15  
; PRIOR APPLICATION NUMBER: GB 0027779.8  
; PRIOR FILING DATE: 2000-11-14  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: Patent In Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 1128  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: coding  
; OTHER INFORMATION: sequence for OmpA-K2S fusion protein  
US-09-987-455-2

Query Match	100.0%;	Score 1128;	DB 11;	Length 1128;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1128;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Oy	1	ATGAAAAGACAGCTATCGGATTCGAGTGGCACTGGCTGGTTTCGTACCGTGGCCAG	60	
Db	1	ATGAAAAGACAGCTATCGGATTCGAGTGGCACTGGCTGGTTTCGTACCGTGGCCAG	60	
Oy	61	CGCGCCTCTGAGGGAACAGTACTGCTACTTTGGGAATGGGTACGCTACCGTGGCCAG	120	





QY 601 TTGAGTCCGAAATACATTGTCATAGGAATTCGATGATGACACATTACGCAATGAC 660  
 Db 601 TTGAGTCCGAAATACATTGTCATAGGAATTCGATGATGACACATTACGCAATGAC 660  
 QY 661 ATTGGCTGCTGACGCTGAATTCGATTCGCTCCGCTGTCGCGCCAGGAGAGAGCTGCTG 720  
 Db 661 ATTGGCTGCTGACGCTGAATTCGATTCGCTCCGCTGTCGCGCCAGGAGAGAGCTGCTG 720  
 QY 721 CGCATGTGTGCTTTCCCGCGGACCTGACGCTGCGGACCTGACGAGTGTGAGCTC 780  
 Db 721 CGCATGTGTGCTTTCCCGCGGACCTGACGCTGCGGACCTGACGAGTGTGAGCTC 780  
 QY 781 TCCGCTAGGCAAGCATGAGGCTTGTCTCTTTCTATTTCGAGCGGCTGAGAGGCT 840  
 Db 781 TCCGCTAGGCAAGCATGAGGCTTGTCTCTTTCTATTTCGAGCGGCTGAGAGGCT 840  
 QY 841 CATGTGACAGTGTACCCATCCAGCGCTGACATCACAACATTTACTTTAAACAGACGTC 900  
 Db 841 CATGTGACAGTGTACCCATCCAGCGCTGACATCACAACATTTACTTTAAACAGACGTC 900  
 QY 901 ACCGCAACATGCTGTGTGCTGGAGACACTCGGAGCGCGGCGCCAGGCAAACTTGCAC 960  
 Db 901 ACCGCAACATGCTGTGTGCTGGAGACACTCGGAGCGCGGCGCCAGGCAAACTTGCAC 960  
 QY 961 GACGCTGCCAGGCGGATTCGGAGGCGCCCTGCTGTCTGAAAGATGGCGCATGACT 1020  
 Db 961 GACGCTGCCAGGCGGATTCGGAGGCGCCCTGCTGTCTGAAAGATGGCGCATGACT 1020  
 QY 1021 TTGGTGGGCATCATCAGCTGGGCTGGGCTGTGGAGAGAGGATGTCGCGGCTGTGAC 1080  
 Db 1021 TTGGTGGGCATCATCAGCTGGGCTGGGCTGTGGAGAGAGGATGTCGCGGCTGTGAC 1080  
 QY 1081 ACAAGGTTACCACTACTAGCTGATTCGTGACAAACATCGGACCG 1128  
 Db 1081 ACAAGGTTACCACTACTAGCTGATTCGTGACAAACATCGGACCG 1128  
 RESULT 3  
 US-09-969-271-6  
 ; Sequence 6, Application US/09969271  
 ; Patent No. US20020098179A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Pfizer Inc. (All designated States except GB and EP (GB));  
 ; APPLICANT: Pfizer Limited (GB and EP (GB) only)  
 ; TITLE OF INVENTION: Pharmaceutical Combinations  
 ; FILE REFERENCE: PCS10951APME  
 ; CURRENT APPLICATION NUMBER: US/09/969,271  
 ; PRIORITY FILING DATE: 2001-10-01  
 ; PRIOR APPLICATION NUMBER: GB 0025473.0  
 ; PRIOR FILING DATE: 2000-10-17  
 ; NUMBER OF SEQ ID NOS: 7  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 6  
 ; LENGTH: 1689  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 US-09-969-271-6  
 Query Match 94.3%; Score 1063.4; DB 9; Length 1689;  
 Best Local Similarity 99.4%; Pred. No. 0;  
 Matches 1067; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
 QY 56 CCCAGGCGGCTCTGAGGAAACAGTACTGCTACTTTGGGAATGGGTCAGCTACCGTG 115  
 Db 614 CCCCTGCTGCTGAGGAAACAGTACTGCTACTTTGGGAATGGGTCAGCTACCGTG 673  
 QY 116 GCAGCAGAGCTCACAGTGGGTGCTCTGCTCCGCTGGAATTCATGATCTGTA 175  
 Db 674 GCAGCAGAGCTCACAGTGGGTGCTCTGCTCCGCTGGAATTCATGATCTGTA 733  
 QY 176 TAGGCAAGTTTACACAGACAGACACAGTCCAGGAGTCCGCTGGGCAACATA 235  
 Db 734 TAGGCAAGTTTACACAGACAGACACAGTCCAGGAGTCCGCTGGGCAACATA 793

QY 236 ATTACTCGCGAATCCTGATGGGATGCCAAGCCCTGCTGCAACGCTGCTGAAGAACCGCA 295  
 Db 794 ATTACTCGCGAATCCTGATGGGATGCCAAGCCCTGCTGCAACGCTGCTGAAGAACCGCA 853  
 QY 296 GGCTGACGCTGGAGTACTGATGATGCTGCTCCTGCTCAACCTGCGGCTTGAGACAGTACA 355  
 Db 854 GGCTGACGCTGGAGTACTGATGATGCTGCTCCTGCTCAACCTGCGGCTTGAGACAGTACA 913  
 QY 356 GCCAGCTCAGCTTTCCGATCAAGAGAGGCTCTTTCGCGACATCGCTCCACCCCTGGC 415  
 Db 914 GCCAGCTCAGCTTTCCGATCAAGAGAGGCTCTTTCGCGACATCGCTCCACCCCTGGC 973  
 QY 416 AGGCTGCCATCTTTGCCAAGCAGAGAGGCTGCCCGGAGAGCGGTTCTGTCGCGGGGCA 475  
 Db 974 AGGCTGCCATCTTTGCCAAGCAGAGAGGCTGCCCGGAGAGCGGTTCTGTCGCGGGGCA 1033  
 QY 476 TACTCATCAGCTCCTGCTGATTTCTCTGCGGCGCACTGCTTCAGAGAGAGTTTCCGC 535  
 Db 1034 TACTCATCAGCTCCTGCTGATTTCTCTGCGGCGCACTGCTTCAGAGAGAGTTTCCGC 1093  
 QY 536 CCCACCACTGACGCTGATCTTGGGCAACATACCGGCTGCTCCCTGCGGAGAGGAGC 595  
 Db 1094 CCCACCACTGACGCTGATCTTGGGCAACATACCGGCTGCTCCCTGCGGAGAGGAGC 1153  
 QY 596 AGAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA 655  
 Db 1154 AGAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA 1213  
 QY 656 ATGACATTTGCGCTGCTGAGCTGAATTCGATTCGCTCCGCTGTCGCGGAGAGAGCG 715  
 Db 1214 ATGACATTTGCGCTGCTGAGCTGAATTCGATTCGCTCCGCTGTCGCGGAGAGAGCG 1273  
 QY 716 TGGTCCGACCTGCTGCTTCCCGCGGAGACCTGACGCTGCGGAGTGGAGAGTGTG 775  
 Db 1274 TGGTCCGACCTGCTGCTTCCCGCGGAGACCTGACGCTGCGGAGTGGAGAGTGTG 1333  
 QY 776 AGCTCTCCGCTACGCAAGCATGAGGCTTGTGCTCTCTTTCTATTTCGAGCGGCTGAAGG 835  
 Db 1334 AGCTCTCCGCTACGCAAGCATGAGGCTTGTGCTCTCTTTCTATTTCGAGCGGCTGAAGG 1393  
 QY 836 AGGCTCATGTGACACTGTATCCCATCCAGCGCTGCAATCAACAATTACTTAAACAGAA 895  
 Db 1394 AGGCTCATGTGACACTGTATCCCATCCAGCGCTGCAATCAACAATTACTTAAACAGAA 1453  
 QY 896 CAGTCACCGCAACATGCTGTGCTGAGACACTCGGAGCGGCGGCGCCAGGCAAACT 955  
 Db 1454 CAGTCACCGCAACATGCTGTGCTGAGACACTCGGAGCGGCGGCGCCAGGCAAACT 1513  
 QY 956 TGCAACGACGCTCCAGGCGGATTCGAGGCGGCTGCTGCTGCTGCTGAACGATGCGCGCA 1015  
 Db 1514 TGCAACGACGCTCCAGGCGGATTCGAGGCGGCTGCTGCTGCTGCTGAACGATGCGCGCA 1573  
 QY 1016 TGACTTTGGTGGGATCATCAGCTGGGCGCTGCGGCTGTCGAGACAGAGATGTCGCGGTG 1075  
 Db 1574 TGACTTTGGTGGGATCATCAGCTGGGCGCTGCGGCTGTCGAGACAGAGATGTCGCGGTG 1633  
 QY 1076 TGTCACAAAGGTTACCACTTACCTAGCTGATTCGCTGACCAACATCGGACCG 1128  
 Db 1634 TGTCACAAAGGTTACCACTTACCTAGCTGATTCGCTGACCAACATCGGACCG 1686

RESULT 4  
 US-10-193-656-7  
 ; Sequence 7, Application US/10193656  
 ; Publication No. US20030096733A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: NY, Tor  
 ; APPLICANT: HOIMDAHL, Rikard  
 ; APPLICANT: LI, Jinan  
 ; TITLE OF INVENTION: NOVEL DRUG TARGETS FOR ARTHRITIS  
 ; FILE REFERENCE: 3810/1J577-US3  
 ; CURRENT APPLICATION NUMBER: US/10/193,656



Db 1050 AGGCTGCAATCTTTCCCAAGCACAGAGGTCGCCGAGAGCGGTCTCTGTCGGGGCA 1109  
Qy 476 TACTCATCAGCTCTCTGCTGGATTCTCTCTGCGGCCACTGCTTCCAGGAGAGTTTCGCC 535  
Db 1110 TACTCATCAGCTCTCTGCTGGATTCTCTCTGCGGCCACTGCTTCCAGGAGAGTTTCGCC 1169  
Qy 536 CCCACCACTGACGTGTATCTTTGGGCAAGAACATACCGGGTGTCTCTGCGAGAGGAGC 595  
Db 1170 CCCACCACTGACGTGTATCTTTGGGCAAGAACATACCGGGTGTCTCTGCGAGAGGAGC 1229  
Qy 596 AGAAATTTGAAGTCGAAATATACATTTCCATAGGAATTCGATGACACTTACGACA 655  
Db 1230 AGAAATTTGAAGTCGAAATATACATTTCCATAGGAATTCGATGACACTTACGACA 1289  
Qy 656 ATGACATTTGCGCTGTCTGACGTGAAATCGGATTCGTCCTGTCGCCAGAGAGCAGCG 715  
Db 1290 ATGACATTTGCGCTGTCTGACGTGAAATCGGATTCGTCCTGTCGCCAGAGAGCAGCG 1349  
Qy 716 TGGTCCGCACTGTGTGCTTCCCGCGGACCTGCACTGCGGACTGGAACGAGTGTG 775  
Db 1350 TGGTCCGCACTGTGTGCTTCCCGCGGACCTGCACTGCGGACTGGAACGAGTGTG 1409  
Qy 776 AGCTCTCGGCTACCGCAAGCATGAGGCTTGTCTCTTTCTATTTCGGAGCGGCTGAAGG 835  
Db 1410 AGCTCTCGGCTACCGCAAGCATGAGGCTTGTCTCTTTCTATTTCGGAGCGGCTGAAGG 1469  
Qy 836 AGGCTCATGTGACAGTGTACCCATCCAGCCGCTGCACATCAACAATTTACTTAACAGAA 895  
Db 1470 AGGCTCATGTGACAGTGTACCCATCCAGCCGCTGCACATCAACAATTTACTTAACAGAA 1529  
Qy 896 CAGTCACCAACAATGTGTGTGTGAGACACTCGGAGCGGGGCCCCAGGCAAACT 955  
Db 1530 CAGTCACCAACAATGTGTGTGTGAGACACTCGGAGCGGGGCCCCAGGCAAACT 1589  
Qy 956 TGCAGAGCGCTGCCAGGCGGATTCGGAGGCGCCCTGCTGTGTCTGAACGATGCGCGCA 1015  
Db 1590 TGCAGAGCGCTGCCAGGCGGATTCGGAGGCGCCCTGCTGTGTCTGAACGATGCGCGCA 1649  
Qy 1016 TGACTTTGTTGGGATCATCAGTCGGGCTGGGCTGTGGAGAGGATGTCCCGGTG 1075  
Db 1650 TGACTTTGTTGGGATCATCAGTCGGGCTGGGCTGTGGAGAGGATGTCCCGGTG 1709  
Qy 1076 TGTACAAAAGTTACCAACTACCTAGACTGGATTCTGTGACAAATGCGACCG 1128  
Db 1710 TGTACAAAAGTTACCAACTACCTAGACTGGATTCTGTGACAAATGCGACCG 1762

RESULT 6

US-09-987-455-4  
; Sequence 4, Application US/09987455  
; Publication No. US20030049729A1  
; GENERAL INFORMATION:  
; APPLICANT: Aranya Manosroi  
; APPLICANT: Jiradej Manosroi  
; APPLICANT: Chatchai Tayapiwatana  
; APPLICANT: Friedrich Goetz  
; APPLICANT: Rolf-Guenther Werner  
; TITLE OF INVENTION: Methods for Large Scale Production of Recombinant  
; FILE OF INVENTION: DNA-derived tpa or K2S Molecules  
; FILE REFERENCE: 0652.2190001  
; CURRENT APPLICATION NUMBER: US/09/987,455  
; CURRENT FILING DATE: 2001-11-14  
; PRIOR APPLICATION NUMBER: 60/268,574  
; PRIOR FILING DATE: 2001-02-15  
; PRIOR APPLICATION NUMBER: GB 002779.8  
; PRIOR FILING DATE: 2000-11-14  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 4  
; LENGTH: 1065  
; TYPE: DNA  
; ORGANISM: Artificial Sequence

FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: coding  
; OTHER INFORMATION: sequence for K2S protein  
US-09-987-455-4

Query Match 94.1%; Score 1062; DB 11; Length 1065;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1062; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 67 TCTCAGGGAACAGTGTCTTCTTTGGGAATGGGTGAGCTTACCGTGGCAGCAGCAGC 126  
Db 1 TCTCAGGGAACAGTGTCTTCTTTGGGAATGGGTGAGCTTACCGTGGCAGCAGCAGC 60  
Qy 127 CTACCCGAGTCGGTGTCTCTCTCCGTTGGAATTCATGATCTCTGATAGGCAAGTT 186  
Db 61 CTACCCGAGTCGGTGTCTCTCTCCGTTGGAATTCATGATCTCTGATAGGCAAGTT 120  
Qy 187 TACACAGACAGACCCAGTGGCCAGGCACTGGGCTGGGCAACATAATTAATCTCGCG 246  
Db 121 TACACAGACAGACCCAGTGGCCAGGCACTGGGCTGGGCAACATAATTAATCTCGCG 180  
Qy 247 AATCTGATGGGATGCCAAGCCCTGTGTCACGCTGTAAGAACCGCAGGCTGACGTGG 306  
Db 181 AATCTGATGGGATGCCAAGCCCTGTGTCACGCTGTAAGAACCGCAGGCTGACGTGG 240  
Qy 307 GAGTACTGTGATGTCCCTCTCTGCTTCCACCTGCGGCTGAGACAGTACAGCAGCCTCAG 366  
Db 241 GAGTACTGTGATGTCCCTCTCTGCTTCCACCTGCGGCTGAGACAGTACAGCAGCCTCAG 300  
Qy 367 TTTCCGATCAAGAGAGGCTCTTCGCGACATGCGCTCCACCCCTGGCAGGCTGCCATC 426  
Db 301 TTTCCGATCAAGAGAGGCTCTTCGCGACATGCGCTCCACCCCTGGCAGGCTGCCATC 360  
Qy 427 TTTCCGAGCAGAGAGTGGCGGAGAGCGGTCTCTGTCGGGGGATACTCATCAGC 486  
Db 361 TTTCCGAGCAGAGAGTGGCGGAGAGCGGTCTCTGTCGGGGGATACTCATCAGC 420  
Qy 487 TCCTGTGGAATCTCTTCCGCGCCACTGCTTCAGGAGAGGTTTCGCGCCCAACCTG 546  
Db 421 TCCTGTGGAATCTCTTCCGCGCCACTGCTTCAGGAGAGGTTTCGCGCCCAACCTG 480  
Qy 547 ACGTGTATTTGGGCAAGAACATACCGGCTGTCTCTGGGAGGAGGAGCAGAAATTTGAA 606  
Db 481 ACGTGTATTTGGGCAAGAACATACCGGCTGTCTCTGGGAGGAGGAGCAGAAATTTGAA 540  
Qy 607 GTCCAAAATAATGTTGTCATAGGAATTCGATGATGACCTTACGACATGACATTCGCG 666  
Db 541 GTCCAAAATAATGTTGTCATAGGAATTCGATGATGACCTTACGACATGACATTCGCG 600  
Qy 667 CTGCTGCAAGTGAATCGGATTCGTCGCTGTGTCGCCAGGAGAGCAGCGTGGTCCGCACT 726  
Db 601 CTGCTGCAAGTGAATCGGATTCGTCGCTGTGTCGCCAGGAGAGCAGCGTGGTCCGCACT 660  
Qy 727 GTGTGCTTCCCGCGGAGCTGAGTCCGAGCTGCGGAGTGGAGTGTGAGCTCTCGGC 786  
Db 661 GTGTGCTTCCCGCGGAGCTGAGTCCGAGCTGCGGAGTGGAGTGTGAGCTCTCGGC 720  
Qy 787 TACGCAAGCATGAGGCTTCTCTCTTCTATTTCGGAGCGGCTGAAGAGGCTCATCTC 846  
Db 721 TACGCAAGCATGAGGCTTCTCTCTTCTATTTCGGAGCGGCTGAAGAGGCTCATCTC 780  
Qy 847 AGACTGTACCCATCCAGCCGCTGCACATCAACAATTTACTTTAAGAACAGTCCCGAC 906  
Db 781 AGACTGTACCCATCCAGCCGCTGCACATCAACAATTTACTTTAAGAACAGTCCCGAC 840  
Qy 907 AACATGTGTGTGTGAGACACTCGGAGCGCGGCGCCAGGCAAACTTGCAGAGCC 966  
Db 841 AACATGTGTGTGTGAGACACTCGGAGCGCGGCGCCAGGCAAACTTGCAGAGCC 900  
Qy 967 TGCCAGGCGGATTCGGGAGGCGCCCTGCTGTGTCGACGATGCCCATGCTTTGTTG 1026  
Db 901 TGCCAGGCGGATTCGGGAGGCGCCCTGCTGTGTCGACGATGCCCATGCTTTGTTG 960

QY 1027 GGCATCATCAGCTGGGGCTTGGCTGTGGACAGAGAGATGTCCCGGTGTGTACACAAAG 1086  
Db |||||  
961 GGCATCATCAGCTGGGGCTTGGCTGTGGACAGAGAGATGTCCCGGTGTGTACACAAAG 1020  
QY 1087 GTTACCAACTACCTAGACTGGATTCGTGACAAACATGCGACCG 1128  
Db |||||

## RESULT 7

US-09-987-455-7  
; Sequence 7, Application US/09987455  
; Publication No. US20030049729A1  
; GENERAL INFORMATION:  
; APPLICANT: Aranya Manosroi  
; APPLICANT: Jiradej Manosroi  
; APPLICANT: Chatchai Tayapiwatana  
; APPLICANT: Friedrich Goetz  
; APPLICANT: Rolf-Guenther Werner  
; TITLE OF INVENTION: Methods for Large Scale Production of Recombinant  
; TITLE OF INVENTION: DNA-Derived tPA or K2S Molecules  
; FILE REFERENCE: 0652.2190001  
; CURRENT APPLICATION NUMBER: US/09/987,455  
; CURRENT FILING DATE: 2001-11-14  
; PRIOR APPLICATION NUMBER: 60/268,574  
; PRIOR FILING DATE: 2001-02-15  
; PRIOR APPLICATION NUMBER: GB 0027779.8  
; PRIOR FILING DATE: 2000-11-14  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 7  
; LENGTH: 1065  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: coding  
; OTHER INFORMATION: sequence for K2S protein  
US-09-987-455-7

Query Match 94.1%; Score 1062; DB 11; Length 1065;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1062; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 67 TCTGAGGAAACAGTACTGTCTTGGGAATGGTCTAGCTACCGTGGCAGCAGCAGC 126  
Db |||||  
1 TCTGAGGAAACAGTACTGTCTTGGGAATGGTCTAGCTACCGTGGCAGCAGCAGC 60  
QY 127 CTCACCGAGTGGGTGCTCTCTGCTCCCGTGGGAATTCATGATCTCATAGGCAAGTT 186  
Db |||||  
61 CTCACCGAGTGGGTGCTCTCTGCTCCCGTGGGAATTCATGATCTCATAGGCAAGTT 120  
QY 187 TACACAGACAGAACCCAGTGGCCAGCAGTGGCCCTGGGCAACATAATCTAGCCGG 246  
Db |||||  
121 TACACAGACAGAACCCAGTGGCCAGCAGTGGCCCTGGGCAACATAATCTAGCCGG 180  
QY 247 AATCTGATGGGATGCCAAGCCCTGTGTGACAGCTGTGGAAGCCGAGCTGACAGTGG 306  
Db |||||  
181 AATCTGATGGGATGCCAAGCCCTGTGTGACAGCTGTGGAAGCCGAGCTGACAGTGG 240  
QY 307 GAGTACTGTGATGTCCTCTCTGCTCCACCTGCGGCTGTGAGACAGTACAGCCAGCTCAG 366  
Db |||||  
241 GAGTACTGTGATGTCCTCTCTGCTCCACCTGCGGCTGTGAGACAGTACAGCCAGCTCAG 300  
QY 367 TTTGCAATCAAGAGGGGCTCTTGGCCAGCATGCTCCACCTGCGGCTGGCAGCTGCCATC 426  
Db |||||  
301 TTTGCAATCAAGAGGGGCTCTTGGCCAGCATGCTCCACCTGCGGCTGGCAGCTGCCATC 360  
QY 427 TTTGCAAGCAGCAGAGAGTGGCCCGGAGAGCGTTCTGTGGGGGGCATACTCATCAGC 486  
Db |||||  
361 TTTGCAAGCAGCAGAGAGTGGCCCGGAGAGCGTTCTGTGGGGGGCATACTCATCAGC 420  
QY 487 TCCTGCTGGATTCCTCTGCGCCGCCCATCTGCTTCAGAGAGAGTTTCCGCCCCACCACTG 546  
Db |||||

Db 421 TCCTGCTGGATTCCTCTGCGCCGCCCATCTGCTTCAGGAGAGGTTTCCGCCCCACCACTG 480  
QY |||||  
547 ACGTGATCTTGGCAGACATACCGGGTGGTCCCTCGCGAGGAGGAGCAGAAATTTGAA 606  
Db |||||  
481 ACGTGATCTTGGCAGACATACCGGGTGGTCCCTCGCGAGGAGGAGCAGAAATTTGAA 540  
QY |||||  
607 GTCGAAAAATACATTTGTCCATAAGGAATTCGATGATGACACTTACGACAATGACATTTGCG 666  
Db |||||  
541 GTCGAAAAATACATTTGTCCATAAGGAATTCGATGATGACACTTACGACAATGACATTTGCG 600  
QY |||||  
667 CTGCTGAGCTGAAATCGGATTTGTCCTCGCTGTGCCAGAGAGCAGGTTGTCGCACT 726  
Db |||||  
601 CTGCTGAGCTGAAATCGGATTTGTCCTCGCTGTGCCAGAGAGCAGGTTGTCGCACT 660  
QY |||||  
727 GTGTGCTTCCCGCGGAGCTGACGTGCGGAGCTGGAGCGAGTGTGAGCTCTCCGCG 786  
Db |||||  
661 GTGTGCTTCCCGCGGAGCTGACGTGCGGAGCTGGAGCGAGTGTGAGCTCTCCGCG 720  
QY |||||  
787 TACGCAAGCATGAGGCTTTGTCTCTCTTCTATTTCGGAGCGGCTGAAGGAGGCTCATGTC 846  
Db |||||  
721 TACGCAAGCATGAGGCTTTGTCTCTCTTCTATTTCGGAGCGGCTGAAGGAGGCTCATGTC 780  
QY |||||  
847 AGACTGTACCATCCAGCGCTGCACATCAACATTTACTTTAACAAGACAGTCAACGAC 906  
Db |||||  
781 AGACTGTACCATCCAGCGCTGCACATCAACATTTACTTTAACAAGACAGTCAACGAC 840  
QY |||||  
907 AACATGCTGTGCTGGAGACACTCGGAGCGGGGCGCCAGGCAAACTTGCACAGCGCC 966  
Db |||||  
841 AACATGCTGTGCTGGAGACACTCGGAGCGGGGCGCCAGGCAAACTTGCACAGCGCC 900  
QY |||||  
967 TGCCAGGCGGATTCGGGAGGCGCCCTGCTGTGTCTGAACGATGCGCGCATGTTGCTG 1026  
Db |||||  
901 TGCCAGGCGGATTCGGGAGGCGCCCTGCTGTGTCTGAACGATGCGCGCATGTTGCTG 960  
QY |||||  
1027 GGCATCATCAGCTGGGGCTTGGCTGTGGACAGAGATGTCCCGGTGTGTACACAAAG 1086  
Db |||||  
961 GGCATCATCAGCTGGGGCTTGGCTGTGGACAGAGATGTCCCGGTGTGTACACAAAG 1020  
QY |||||  
1087 GTTACCAACTACCTAGACTGGATTCGTGACAAACATGCGACCG 1128  
Db |||||  
1021 GTTACCAACTACCTAGACTGGATTCGTGACAAACATGCGACCG 1062

## RESULT 8

US-09-974-298-144  
; Sequence 144, Application US/09974298  
; Patent No. US20020156263A1  
; GENERAL INFORMATION:  
; APPLICANT: Chen, Huel-Mei  
; TITLE OF INVENTION: GENES EXPRESSED IN BREAST CANCER  
; FILE REFERENCE: PA-0037 P  
; CURRENT APPLICATION NUMBER: US/09/974,298  
; CURRENT FILING DATE: 2001-10-04  
; PRIOR APPLICATION NUMBER: 60/238,331  
; PRIOR FILING DATE: 2000-05-10  
; NUMBER OF SEQ ID NOS: 194  
; SOFTWARE: PERL Program  
; SEQ ID NO 144  
; LENGTH: 2641  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; OTHER INFORMATION: Incyte ID No. US20020156263A1 1001470CB1  
; NAME/KEY: unsure  
; LOCATION: 2635  
; OTHER INFORMATION: a, t, c, g, or other  
US-09-974-298-144

Query Match 94.1%; Score 1061.8; DB 10; Length 2641;  
Best Local Similarity 99.3%; Pred. No. 0;  
Matches 1066; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

```
Qy 56 CCCAGCGGCTCTGAGGAAACAGTGTACTTTGGGAATGGGTGACGCTACCGTG 115
Db |||
Qy 810 CCCCTGCGTCTGAGGAAACAGTGTACTTTGGGAATGGGTGACGCTACCGTG 869
Db |||
Qy 116 GCAGCAGAGCTCACCGAGTGGGTGCTCTGCTCCGCTCGGTAATTCATGATCCTGA 175
Db |||
Qy 870 GCAGCAGAGCTCACCGAGTGGGTGCTCTGCTCCGCTCGGTAATTCATGATCCTGA 929
Db |||
Qy 176 TAGCAAGGTTTACACAGCAGAACCCAGTGGCCAGGCACTGGGCTGGGCAACATA 235
Db |||
Qy 930 TAGCAAGGTTTACACAGCAGAACCCAGTGGCCAGGCACTGGGCTGGGCAACATA 989
Db |||
Qy 216 ATTACTGCGGAATCTGATGGGATGCAAGCCTGTGTGCACTGTGTGAAGACCGCA 295
Db |||
Qy 990 ATTACTGCGGAATCTGATGGGATGCAAGCCTGTGTGCACTGTGTGAAGACCGCA 1049
Db |||
Qy 296 GGCTGACGTGGGAGTGTGATGCGCTCTGCTCCACCTGCGGCTGAGACAGTACA 355
Db |||
Qy 1050 GGCTGACGTGGGAGTGTGATGCGCTCTGCTCCACCTGCGGCTGAGACAGTACA 1109
Db |||
Qy 356 GCCAGCCTCAGTTTCGATCAAAAGGAGGCTCTTGGCGGACATCGCCTCCACCCCTGGC 415
Db |||
Qy 1110 GCCAGCCTCAGTTTCGATCAAAAGGAGGCTCTTGGCGGACATCGCCTCCACCCCTGGC 1169
Db |||
Qy 416 AGGCTGCCATCTTTGCCAAGCACAGGAGGTGCGCCGAGAGCGGTTCTGTGCGGGGCA 475
Db |||
Qy 1170 AGGCTGCCATCTTTGCCAAGCACAGGAGGTGCGCCGAGAGCGGTTCTGTGCGGGGCA 1229
Db |||
Qy 476 TACTCATCAGCTCCTGCTGGATTCTCTGCGGCCACTGCTCCAGGAGGTTTCGC 535
Db |||
Qy 1230 TACTCATCAGCTCCTGCTGGATTCTCTGCGGCCACTGCTCCAGGAGGTTTCGC 1289
Db |||
Qy 536 CCCACCCTGACGCTGATCTTTGGGCAAGACATACCGGGTGTCTCTGGCGAGGAGG 595
Db |||
Qy 1290 CCCACCCTGACGCTGATCTTTGGGCAAGACATACCGGGTGTCTCTGGCGAGGAGG 1349
Db |||
Qy 596 AGAAATTTGAAGTCGAAATATACATTTGCCAATAGGAATTCATGATGACACTTACGACA 655
Db |||
Qy 1350 AGAAATTTGAAGTCGAAATATACATTTGCCAATAGGAATTCATGATGACACTTACGACA 1409
Db |||
Qy 656 ATGACATTTGCTGCTGAGTGAATTCGATTCGCTCCGCTGTGCCAGGAGCAGCG 715
Db |||
Qy 1410 ATGACATTTGCTGCTGAGTGAATTCGATTCGCTCCGCTGTGCCAGGAGCAGCG 1469
Db |||
Qy 716 TGGTCCGCACTGTGTGCTTCCCGCGGACCTGACGCTCCGCTGTGACGAGTGTG 775
Db |||
Qy 1470 TGGTCCGCACTGTGTGCTTCCCGCGGACCTGACGCTCCGCTGTGACGAGTGTG 1529
Db |||
Qy 776 AGCTCTCCGCTACGGCAAGCATGAGGCTTGTCTCTTTCTATTCGAGCGGCTGAAG 835
Db |||
Qy 1530 AGCTCTCCGCTACGGCAAGCATGAGGCTTGTCTCTTTCTATTCGAGCGGCTGAAG 1589
Db |||
Qy 836 AGGCTCATGTGAGTGTACCATTCAGCCGCTGACATCAACATTTACTTACAGAA 895
Db |||
Qy 1590 AGGCTCATGTGAGTGTACCATTCAGCCGCTGACATCAACATTTACTTACAGAA 1649
Db |||
Qy 896 CAGTCACCGCAACATGTGTGTGAGACACTCGGAGCGGGGCCCCAGAGCAAACT 955
Db |||
Qy 1650 CAGTCACCGCAACATGTGTGTGAGACACTCGGAGCGGGGCCCCAGAGCAAACT 1709
Db |||
Qy 956 TGCACGACGCTGCCAGGCGATTTCGGAGCGCCCCCTGCTGTGTCTGAACGATGGCGCA 1015
Db |||
Qy 1710 TGCACGACGCTGCCAGGCGATTTCGGAGCGCCCCCTGCTGTGTCTGAACGATGGCGCA 1769
Db |||
Qy 1016 TGACTTTGGTGGGATCATAGCTGGGCTGTGGCTGTGGACAGAGATGTCCGGGTG 1075
Db |||
Qy 1770 TGACTTTGGTGGGATCATAGCTGGGCTGTGGCTGTGGACAGAGATGTCCGGGTG 1829
Db |||
Qy 1076 TGTACACAAAGGTTACCAACTTACCTAGACTGGATTTCGTGACAAATGCGACCG 1128
Db |||
Qy 1830 TGTACACAAAGGTTACCAACTTACCTAGACTGGATTTCGTGACAAATGCGACCG 1882
Db |||
```

RESULT 9

```
US-10-007-926A-433
; Sequence 433, Application US/10007926A
; Publication No. US20030143539A1
; GENERAL INFORMATION:
; APPLICANT: BERTUCCI, FRANCOIS
; APPLICANT: HOULGATTE, REMI
; APPLICANT: BIRNBAUM, DANIEL
; APPLICANT: NGUYEN, CATHERINE
; APPLICANT: VIENS, PATRICE
; APPLICANT: FERT, VINCENT
; TITLE OF INVENTION: GENE EXPRESSION PROFILING OF PRIMARY BREAST CARCINOMAS
; FILE REFERENCE: 1546-R-00
; CURRENT APPLICATION NUMBER: US/10/007,926A
; CURRENT FILING DATE: 2001-12-07
; PRIOR APPLICATION NUMBER: 60/254,090
; PRIOR FILING DATE: 2000-12-08
; NUMBER OF SEQ ID NOS: 468
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 433
; LENGTH: 329
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: 5' terminal sequence. Homo sapiens plasminogen activator (PLAT)
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (57)
; OTHER INFORMATION: a, t, c or g
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (82)
; OTHER INFORMATION: a, t, c or g
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (159)
; OTHER INFORMATION: a, t, c or g
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (264)
; OTHER INFORMATION: a, t, c or g
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (278)
; OTHER INFORMATION: a, t, c or g
US-10-007-926A-433
```

Query Match 20.1%; Score 227.2; DB 12; Length 329;  
Best Local Similarity 97.4%; Pred. No. 5.9e-62;  
Matches 229; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```
Qy 894 AACAGTCACGCAACATGCTGTGTGAGAGACACTCGGAGCGGGGCCAGGCAAA 953
Db 1 AACAGTCACGCAACATGCTGTGTGAGAGACACTCGGAGCGGGGCCAGGCAAA 60
Qy 954 CTTGACAGCGCTTCGAGGCGGATTCGGGAGGCCCTGTGTGTCTGAACGATGCGC 1013
Db 61 CTTGACAGCGCTTCGAGGCGGATTCGGGAGGCCCTGTGTGTCTGAACGATGCGC 120
Qy 1014 CATGACTTTTGGTGGGCATCATCAGCTGGGCGCTGTGAGCAGAGGATGTCCCGGG 1073
Db 121 CATGACTTTTGGTGGGCATCATCAGCTGGGCGCTGTGAGCAGAGGATGTCCCGGG 180
Qy 1074 TGTGTACCAAGGTTACCAACTACCTAGACTGTGATTCGTGACAAATGCGACCG 1128
Db 181 TGTGTACCAAGGTTACCAACTACCTAGACTGTGATTCGTGACAAATGCGACCG 235
```

RESULT 10

```
US-09-880-503-15
; Sequence 15, Application US/09880503
; Patent No. US20020131964A1
```

Qy	843	TGTCAGACTGTATCCCATTCAGCGGCTGCAATCAAAATTTACTTTAAAGAACATGCAC	902
Db	924	TGTGAAGCTGATTTTCCACCGGGAGTGTGAGCAGGCCCACTACTACGGCTCTGAAGTCA	983
Qy	903	CGAACACATGCTGTGTGCTGGAGACACTCGGAGCGGGCCCCAGGCCAACTTGACCA	962
Db	984	CACCAAAATGCTATGTGCTGCTGACCCCC-----AATGGAAAAACAGA	1025
Qy	963	CGCTCCAGGGCGGATTCGGGAGGCCCTCGTGTGTCTGAACGATGGCCGCGATGACTTT	1022
Db	1026	TTCTCTGCCAGGAGACTCAGGGGACCCCTGCTGTCTCCCTCCAAGCGCGCATGACTTT	1085
Qy	1023	GTGGGCATCATCAGCTGGGGCTTGGGCTGTGGACAGAAAGGATGTCCCGGTGTGTACAC	1082
Db	1086	GACTCGAAATGTGTGAGCTGGGGCGGTGATGTGCTCTGAAGGACAAGCCAGGCGCTTACAC	1145
Qy	1083	AAAGGTTACCACTACTAGACTGGATTGCG	1112
Db	1146	GAGAGTCTCACACTTCTTACCCTGGATCGG	1175
RESULT 11			
US-10-198-846-12748			
; Sequence 12748, Application US/10198846			
; Publication No. US2003009974A1			
; GENERAL INFORMATION:			
; APPLICANT: Lillie, James			
; APPLICANT: Xu, Yongyao			
; APPLICANT: Wang, Youzhen			
; APPLICANT: Steinmann, Kathleen			
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS			
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND			
; TITLE OF INVENTION: THERAPY OF BREAST CANCER			
; FILE REFERENCE: MRI-049			
; CURRENT APPLICATION NUMBER: US/10/198,846			
; CURRENT FILING DATE: 2002-07-18			
; PRIOR APPLICATION NUMBER: 60/306,220			
; PRIOR FILING DATE: 2001-07-18			
; NUMBER OF SEQ ID NOS: 14084			
; SOFTWARE: FastSeq for Windows Version 4.0			
; SEQ ID NO 12748			
; LENGTH: 1415			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
; FEATURE:			
; NAME/KEY: misc_feature			
; LOCATION: 1, 2, 3, 1414, 1415			
; OTHER INFORMATION: n = A,T,C or G			
US-10-198-846-12748			
Query Match 16.7%; Score 188; DB 14; Length 1415;			
Best Local Similarity 91.1%; Pred. No. 3.1e-49;			
Matches 255; Conservative 0; Mismatches 20; Indels 5; Gaps 5;			
Qy	56	CCACGCGCGCTCTGAGGAAACAGTACTGTCTTTGGGAATGGTCAAGCTACCGTG	115
Db	895	CCCTCGCTGCTCTGAGGAAACAGTACTGT-TACTTTGGGAATGGTCAAGCTACCGTG	953
Qy	116	GCACGCACACGCTTCAACGAGTCGGGTGCCTCT-GCCTCCCGTGGAAATTCATGATCCTG	174
Db	954	GCACGCACAG-CTCACCGAGTCGGGTGCCTCTCTGGCCCTCCCGTGGAAATTCATGATCCTG	1012
Qy	175	ATAGGCAAGTTTACACAGCAGAACCCAGTCGCCAGGCACTGGGGCTGGGCAACAT	234
Db	1013	ATAGGCAAGTTTACACAGCAGAACCCAGTCGCCAGGCACTGGGGCTGGGCAACAT	1072
Qy	235	AATTACTGCCGGAATCCTCATGGGAATGCCAAGCCCTGTGTGCACCTGTCTGAA-GAACCG	293
Db	1073	AAATTAATCCGAAT-CTGATGGGAATGCCAAGCCCTGTGTGCACCTGTCTGAAACCG	1131
Qy	294	CAGGCTGAGCTGGGAGTACTGTGATGTGCCCTCTGCTCC	333

Db 1132 CAGGCTGACGTGGGAGTACTGTGATGTGGCCTCTGTCAACC 1171

RESULT 12

US-09-735-705-122  
; Sequence 122, Application US/09735705  
; Patent No. US20020052329A1

GENERAL INFORMATION:  
; APPLICANT: Wang, Tongtong

; APPLICANT: Fan, Liqun

; APPLICANT: Kalos, Michael D.

; APPLICANT: Bangur, Chaitanya S.

; APPLICANT: Hosken, Nancy

; APPLICANT: Fanger, Gary R.

; APPLICANT: Li, Samuel X.

; APPLICANT: Wang, Aijun

; APPLICANT: Skeiky, Yasir A.W.

; APPLICANT: Henderson, Robert A.

; APPLICANT: McNeill, Patricia D.

; APPLICANT: Fanger, Neil

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY

; FILE REFERENCE: 210121.455C14

; CURRENT APPLICATION NUMBER: US/09/735,705

; CURRENT FILING DATE: 2000-12-12

; NUMBER OF SEQ ID NOS: 419

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 122

; LENGTH: 1475

; TYPE: DNA

; ORGANISM: Homo sapien

US-09-735-705-122

Query Match 15.1%; Score 170.8; DB 9; Length 1475;

Best Local Similarity 51.3%; Pred. No. 9.6e-44;

Matches 551; Conservative 0; Mismatches 457; Indels 66; Gaps 4;

Qy 84 CTGCTACTTTGGGAATGGGTGAGCTACCGTGGCAACGACGCTCAACGAGTGGGTGC 143  
Db 287 CTGCTATGAGGGAATGGTCACTTTACCGAGGAAGGCCAGCACTGACACCATGGGCG 346  
Qy 144 CTCTGCTCCCGTGGATTCATGATCTCTATAGGCAAGTTTACACAGACAGAACCC 203  
Db 347 GCGTGGCTGGCTGGGAACTGCGCACTGCTTTCAGCAAACTGACATGCGCCAGATC 406  
Qy 204 CAGTGCCAGGACCTGGGCTGGGCAACATAATTTACTGCGGAACTCTGATGGGATGC 263  
Db 407 TGATGCTCTTTCAGTGGGCTGGGAAACATAATTTACTGCGGAACTCTGATGGGATGC 466  
Qy 264 CAAGCCCTGGTGCCAGTGTGAAGAACCGGAGGCTGACGTGGGAGTACTGTGATGTC 323  
Db 467 GCGACCTCTGTGTATGTGCAAGTGGGCTAAAGCCGCTTGTCCAGAGTGCATGGTGA 526  
Qy 324 CTCCTGCTCCACCTGGGCTGGAG----- 347  
Db 527 TGACTGGCAGATGGAAAAAGCCCTCTCTCTCCAGAGAAATTAATAATTCAGTGTGG 586  
Qy 348 ---ACAGTACAGCAGCTCAGTTTTCGATCAAGAGGAGGCTCTTCGCGGACATCGCTC 404  
Db 587 CCAAAAGACTCTGAGGCCCCCTTTAGATTTATGGGGGAGAAATTCACACCATCGAGAA 646  
Qy 405 CCACCCCTGGCAGGCTGCATCTTTTGGCAACAGAGAGGTGCGCCGCGAGAGCGGTTCCT 464  
Db 647 CCAGCCCTGGTTTGGGCGCACTACAGGAGCAACCGGG---GGGCTCTGTCACTAGT 703  
Qy 465 GTGGGGGGGATACATCAGCTCTGTGATTTCTCTCTGCGCGCCACATGCTTCCAGGA 524  
Db 704 GTGTGGAGGAGCTCATCAGCCCTTGTCTGGGTGATCAGCGCCACACACTGCTTCAITGA 763  
Qy 525 GAGGTTTTCCGCCCCACACCTGACGGTGTATCTTGGGAGAACTACCGGGTGGTCCCTGG 584  
Db 764 TTACCCAAAGAGGAGGACTACATGCTACTCTGGGTGCTCAAGGCTTAACCTCAACAC 823

Qy 585 CGAGGAGGAGCAGAAATTTGAAGTCGAAAAATAATTTGTCTCATAGGAATTCGATGATGA 644  
Db 824 GCAAGGGGAGATGAAGTTTGAAGTGGAAAACTCTCTACACAGGACTACAGCGCTGA 883  
Qy 645 CA-----CTTACGACATGACATTTGGCTGCTGACGTGAAATCGGATTCGTCGCGTG 698  
Db 884 CAGCTTGTCTACACACAGCATTGCTTGTCTGCTGAGATTCGTTTCCAAAGGAGGAGGTG 943  
Qy 699 TGCCAGGAGAGCAGCGTGTCCGCACTGTGTGCTTCCCGCGGCGACCTGCGAGCTGCC 758  
Db 944 TGCGACCCATCCCGACTATACAGACCACTGCTGCTGCTCGATGTATACGATCCCCA 1003  
Qy 759 GGACTGAGCGAGTGTGAGCTCTCCGCTACGGCAAGCATGAGGCTTTGCTCTTCTTA 818  
Db 1004 GTTTGGCACAAGCTGTGAGATCACTGGCTTTTGGAAAAAGAAATTTCTACCGACTATCTTA 1063  
Qy 819 TTGGAGCGGCTGAAGGAGGCTCATGTGACACTGTATCCCATCCAGCGCGCTGCACATACA 878  
Db 1064 TCGGAGCAGCTGAAGATGACTGTTGTGAAGTGAATTTCCACCGGAGTGTACAGAGCC 1123  
Qy 879 ACATTTACTTTAAACAGAACAGTCAACCAACATGCTGTGTCTGGAGACACTCCGAGCGG 938  
Db 1124 CCCTACTACGGCTCTGAAGTCAACCAACATGCTGTGTCTGCTGCTGCTGCTGCTGCTG 1180  
Qy 939 CGGCGCCCGAGCAAACTTTGCAAGCGCTGCGGAGCGATTTGGGAGGCGCCCTGGTGTG 998  
Db 1181 -----GAAACACAGATTTCTGCGAGGAGACTCAGGGGGGACCCCTCGTCTG 1225  
Qy 999 TCTGAACGATGGCGCATGACTTTGGTGGGCATCATCAGCTGGGCGCTGGGCTGTGGACA 1058  
Db 1226 TTCCCTCAAGGCGCGCATGACTTTGACTGGAATTTGAGCTGGGGCGCTGGAGTGGCCCT 1285  
Qy 1059 GAAGGATGTCCTGGGTGTGTACACAAAGGTTTACCACTACCTAGACTGGATTCG 1112  
Db 1286 GAAGGACAAAGCGGCTGTACAGGAGTGTACACTTTTACCTTGTGATTCG 1339

RESULT 13

US-09-850-716A-122  
; Sequence 122, Application US/09850716A  
; Patent No. US20020115139A1

GENERAL INFORMATION:  
; APPLICANT: Kalos, Michael D.

; APPLICANT: McNeill, Patricia D.

; APPLICANT: Retter, Marc W.

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY

; FILE REFERENCE: 210121.455C15

; CURRENT APPLICATION NUMBER: US/09/850,716A

; CURRENT FILING DATE: 2001-05-07

; NUMBER OF SEQ ID NOS: 440

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 122

; LENGTH: 1475

; TYPE: DNA

; ORGANISM: Homo sapien

US-09-850-716A-122

Query Match 15.1%; Score 170.8; DB 10; Length 1475;

Best Local Similarity 51.3%; Pred. No. 9.6e-44;

Matches 551; Conservative 0; Mismatches 457; Indels 66; Gaps 4;

Qy 84 CTGCTACTTTGGGAATGGGTGAGCTACCGTGGCAGCAGCAGCTCAACGAGTGGGTGC 143  
Db 287 CTGCTATGAGGGAATGGTCACTTTTACCGAGGAAGGCCAGCACTGACACCATGGGCG 346  
Qy 144 CTCTGCTCCCGTGGAAATTCATGATCTCTGATAGGCAAGTTTACACAGCAGAACACC 203  
Db 347 GCGTGGCTGGCTGGGAACTGCTGCACTGTCTTTCAGCAAACTGATCCATGCCACAGATC 406  
Qy 204 CAGTGCCAGGACCTGGGCGCTGGGCAACATAATTTACTGCGGAACTCTGATGGGATGC 263  
Db 407 TGATGCTCTTTCAGTGGGCTGGGAAACATAATTTACTGAGGAACCCAGACACCGGAG 466







Query Match	15.1%;	Score 170.8;	DB 11;	Length 1475;
Best Local Similarity	51.3%;	Pred. No. 9.6e-44;		
Matches 551;	Conservative 0;	Mismatches 457;	Indels 66;	Gaps 4;
QY	84	CTGCTACTTTGGGAATGGGTGAGCTACCGTGGGACGCACACAGCTCACCGAGTCGGGTGC	143	
Db	287	CTGCTATGAGGGGNAUGTCACTTTACGAGGAAAGCCACGACTGACACCATGGCCGC	346	
QY	144	CTCTGCTCCCTGGGAATTCATGATCTGTATAGGCAAGGTTTACACGACACAGAACCC	203	
Db	347	GCCCTGCTGCCCTGGAACTCTGCCACTGCTCTTCAGCAAAACGTACCATGCCACAGATC	406	
QY	204	CAGTGGCCAGGCACCTGGGCTGGGCAACAATAATTAATCTGCGGNAATCTGATGGGATGC	263	
Db	407	TGATGCTCTTTTCAGCTGGGCTGGGGAAACAATAATTAATCTGCAGGAACCTCAGACAACCGGAG	466	
QY	264	CAAGCCCTGGTGCCACGTGTGTGAAGAACCGCAGGCTCAAGCTGGGAGTACTGTGATGTGCC	323	
Db	467	GCGNCCCTGGTGCTATGTGACGTGGGCTTAAGCCGCTTGTCCAGAGTGCATGTGTGCA	526	
QY	324	CTCTGTCTCCACTGCGGCTGTAG-----	347	-----
Db	527	TGACTGGCCAGATGGAATAAAAGCCCTCTCTCTCTCCAGAAAGATTAAAAATTCAGTGTGG	586	
QY	348	---ACAGTACAGCCAGCCTCAGTTTCGCATCAAGAGGGGCTCTTGCCGCACATCGCCTC	404	
Db	587	CCAAAAGACTCTGAGGCCCGCGCTTAAGATTATTGGGGGAGAAATTCACCAACCATTCGAGAA	646	
QY	405	CCACCCCTGGCAGGCTGCCATCTTTGCAAGCACAGGAGGTGCGCCCGGAGAGCGGTTCCT	464	
Db	647	CCAGCCCTGGTTTGGCGCCATCTACAGGAGGCACCGGG---GGGCTCTGTCACTACGT	703	
QY	465	GTGGGGGGCATACTCATAGTCTCTGCTGGATTCTCTCTGCGCGCCCACTGCTTCCAGGA	524	

Search completed: August 19, 2003, 14:22:27  
Job time : 320.307 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 18, 2003, 22:44:33 ; Search time 3299.95 Seconds  
(without alignments)  
8307.845 Million cell updates/sec

Title: US-09-987-455-5

Perfect score: 1128

Sequence: 1 atgaaaagacagctatgc.....ttctgacaaatgcgacgc 1128

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*

- 1: em\_estba:\*
- 2: em\_esthum:\*
- 3: em\_estin:\*
- 4: em\_estmu:\*
- 5: em\_estov:\*
- 6: em\_estpl:\*
- 7: em\_estro:\*
- 8: em\_htc:\*
- 9: gb\_est1:\*
- 10: gb\_est2:\*
- 11: gb\_htc:\*
- 12: gb\_est3:\*
- 13: gb\_est4:\*
- 14: gb\_est5:\*
- 15: em\_estfun:\*
- 16: em\_estom:\*
- 17: em\_gss\_hum:\*
- 18: em\_gss\_inv:\*
- 19: em\_gss\_pln:\*
- 20: em\_gss\_vrt:\*
- 21: em\_gss\_fun:\*
- 22: em\_gss\_mam:\*
- 23: em\_gss\_mus:\*
- 24: em\_gss\_pro:\*
- 25: em\_gss\_rod:\*
- 26: em\_gss\_phg:\*
- 27: em\_gss\_vrl:\*
- 28: gb\_gss1:\*
- 29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
c 1	895	79.3	930	13	BX329047
2	853	75.6	912	13	BUI49958
3	850.2	75.4	962	13	BX389609
4	818.6	72.6	891	13	BUI46167

5	813.8	72.1	993	13	BQ278177
6	811.2	71.9	883	13	BQ690978
7	807.6	71.6	918	13	BX389608
8	796.2	70.6	930	13	BUI45014
9	782.8	69.4	888	13	BUI57720
10	765.2	67.8	950	13	BQ927899
11	764.4	67.8	916	12	BI765247
12	706.8	62.7	948	10	BE616613
13	700.4	62.1	958	12	BG763582
14	694.4	61.6	974	13	BQ960216
15	693.4	61.5	709	10	BE732704
16	688.2	61.0	1060	12	BM556999
17	677	60.0	1042	10	BE732413
18	672	59.6	957	13	BQ687538
19	665.8	59.0	948	13	BQ687779
20	665.2	59.0	704	9	AUI134301
21	655.4	58.1	818	9	AUI124602
22	655.2	58.1	903	13	BQ689095
23	653.2	57.9	830	13	BQ896938
24	653.2	57.9	867	13	BUI79903
25	652.2	57.8	904	13	BQ689840
26	652.2	57.8	911	13	BQ684734
27	651	57.7	818	12	BM006474
28	649.6	57.6	797	13	BQ879911
29	642.2	56.9	1037	13	BUI57493
30	631.2	56.0	813	10	BG120780
31	629.8	55.8	663	9	AUI134463
32	624	55.3	624	10	BE549489
33	585.4	51.9	587	10	BE394944
34	561	49.7	561	9	AL047284
35	553.4	49.1	555	14	CB128855
36	535.4	47.5	537	14	CB155052
37	534	47.3	964	12	BG769384
38	520.4	46.1	905	13	BUI859271
39	514.8	45.6	883	14	CD359959
40	514.4	45.6	516	14	CB140923
41	493.4	43.7	916	13	BUI57241
42	487.8	43.2	527	9	AW630119
43	482.8	42.8	553	9	AW394127
44	474.8	42.1	917	13	BUI90904
45	472.8	41.9	937	13	BQ679501

ALIGNMENTS

RESULT 1  
BX329047/c  
LOCUS  
DEFINITION BX329047 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA  
Clone CSODI039YK15 3-PRIME, mRNA sequence.  
ACCESSION BX329047  
VERSION BX329047.1 GI:30307809  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 930)  
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.  
TITLE Full-length cDNA libraries and normalization  
JOURNAL Unpublished  
COMMENT Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 EVRY cedex - France  
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
Library was constructed by Life Technologies, a division of  
Invitrogen. This sequence belongs to sequence cluster 4498.r For  
more information about this cluster, see  
http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CS0BAI0212D08\_C  
S019801&cluster=4498.r. Contact : Feng Liang Email :  
fliang@lifetech.com URL : http://fulllength.invitrogen.com/  
InvitroGen Corporation 1600 Faraday Avenue Genoscope sequence ID :

```

CS0BAI0212D08 CS01980 1.
FEATURES
  Location/Qualifiers
    1..930
      /organism="Homo sapiens"
      /mol_type="mRNA"
      /db_xref="taxon:9606"
      /clone="CS0D1039YK15"
      /tissue_type="PLACENTA COT 25-NORMALIZED"
      /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
      /note="1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT      184 a   268 c   282 g   194 t      2 others
ORIGIN
Query Match      79.3%; Score 895; DB 13; Length 930;
Best Local Similarity 98.6%; Pred. No. 2.9e-22;
Matches 912; Conservative 0; Mismatches 12; Indels 1; Gaps 1;
QY 67 TCTGAGGAAACAGTGACTGCTTCTGGGAATGGTTCAGCTACCGTGGCAGCCAGC 126
DB 927 TTGTGGGAAACAGTGACTGCTTAC-TTGGGAATGGTTCAGCTACCGTGGCAGCCAGC 869
QY 127 CTCACCGAGTCGGGTGCTCTGCTCCCGTGGGAATTCATGATCCTGATAGGCAAGTT 186
DB 868 CTCACCGAGTCGGGTGCTCTGCTCCCGTGGGAATTCATGATCCTGATAGGCAAGTT 809
QY 187 TACACAGACAGAACCCAGTCGCCAGCAGTCGGCTGGGCAACATTAATTAATTCGCGG 246
DB 808 TACACAGACAGAACCCAGTCGCCAGCAGTCGGCTGGGCAACATTAATTAATTCGCGG 749
QY 247 AATCTGATGGGATGCCAAGCCCTGGTGCACCTGCTGAGAACCCGAGCTGACCTGG 306
DB 748 AATCTGATGGGATGCCAAGCCCTGGTGCACCTGCTGAGAACCCGAGCTGACCTGG 689
QY 307 GAGTACTGTGATGTCCTCTGCTCCACCTGCGGCTGAGACAGTACAGCAGCAGCTCAG 366
DB 688 GAGTACTGTGATGTCCTCTGCTCCACCTGCGGCTGAGACAGTACAGCAGCAGCTCAG 629
QY 367 TTTCCATCAAGAGGAGGCTTTTGGCGAATCGCTCCACCCCTGCGAGGCTGCCATC 426
DB 628 TTTCCATCAAGAGGAGGCTTTTGGCGAATCGCTCCACCCCTGCGAGGCTGCCATC 569
QY 427 TTTCCAGACAGAGGCTTCCCGGAGAGGCTTCTGTCGGGGGATACATCATCAGC 486
DB 568 TTTCCAGACAGAGGCTTCCCGGAGAGGCTTCTGTCGGGGGATACATCATCAGC 509
QY 487 TCCTGCTGGATTCCTCTGCGCGCCACCTGCTTCCAGAGAGGTTTCCGCCCCACACCTG 546
DB 508 TCCTGCTGGATTCCTCTGCGCGCCACCTGCTTCCAGAGAGGTTTCCGCCCCACACCTG 449
QY 547 ACGGTGATCTTTGGGAGAACATACCGGTGTCTGTCGGAGGAGGAGACGAAATTTGAA 606
DB 448 ACGGTGATCTTTGGGAGAACATACCGGTGTCTGTCGGAGGAGGAGACGAAATTTGAA 389
QY 607 GTCGAAATATCATGTCATTAAGGAATTCATGATGACATTCAGCAATGACATTCGC 666
DB 388 GTCGAAATATCATGTCATTAAGGAATTCATGATGACATTCAGCAATGACATTCGC 329
QY 667 CTGCTGACAGCTGAATCCGATTCGTCCCGCTGTGCCAGAGAGCAGGTCGTCCGCACT 726
DB 328 CTGCTGACAGCTGAATCCGATTCGTCCCGCTGTGCCAGAGAGCAGGTCGTCCGCACT 269
QY 727 GTGTGCTTCCCGCGCGGACCTGAGCTGCGGACCTGAGAGGAGTGTGAGCTTCCGGC 786
DB 268 GTGTGCTTCCCGCGCGGACCTGAGCTGCGGACCTGAGAGGAGTGTGAGCTTCCGGC 209
QY 787 TACGCGACATGAGGCTTCTCTCTTTTATTCGGAGCGGCTGAGAGGCTCATGTC 846
DB 208 TACGCGACATGAGGCTTCTCTCTTTTATTCGGAGCGGCTGAGAGGCTCATGTC 149
QY 847 AGACTGTACCCATCCAGCGGCTGCACATCAACAATTTTACTTAACAGAAAGTCAACGAC 906
148 AGACTGTACCCATCCAGCGGCTGCACATCAACAATTTTACTTAACAGAAAGTCAACGAC 89
907 AACATGCTGTGTCTGGAGACACTCGAGAGCGGGGGCCCGAGCAAACTTGCACAGCC 966
88 AACATGCTGTGTCTGGAGACACTCGAGAGCGGGGGCCCGAGCAAACTTGCACAGCC 29
967 TGCCAGGGCGATTTCGGAGGCGCC 991
28 TGCCAGGGCGATTTCGGAGGCGCC 4
RESULT 2
LOCUS      BUI49958      912 bp      mRNA      linear      EST 03-SEP-2002
DEFINITION AGENCOURT_8137213 lupski_dorsal_root_ganglion Homo sapiens cDNA
ACCESSION   BUI49958
VERSION     BUI49958.1 GI:22663490
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 912)
AUTHORS    NIH-MGC http://mgi.nci.nih.gov/
TITLE      National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL     Unpublished
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: rgs@nih.gov
            Tissue Procurement: Dr. James R. Lupski
            CDNA Library Preparation: Life Technologies, Inc.
            CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LIA3572 row: m column: 16
            High quality sequence start: 18
            High quality sequence stop: 644.
FEATURES
  source
    1..912
      /organism="Homo sapiens"
      /mol_type="mRNA"
      /db_xref="taxon:9606"
      /clone="IMAGE:6184119"
      /sex="male"
      /tissue_type="dorsal root ganglia"
      /dev_stage="adult, 36 yr"
      /lab_host="DH10B"
      /clone_lib="Lupski dorsal root ganglion"
      /notes="vector: pCMV-SPORT6 (Life Technologies); Site_1:
            NotI; Site_2: SalI; cDNA made by oligo-dT priming.
            Directionally cloned using the following adaptors:
            5'-TCGACCCAGCGCTCCG-3' and
            5'-GAGTACTTCTAGATCGGAGCGGCGCT(15)-3'. Size selected >
            1 kb for average insert length 1.7 kb. This is a primary
            library, non-amplified. Library constructed by Life
            Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
            College of Medicine) and is available through Life
            Technologies."
BASE COUNT      189 a   272 c   271 g   179 t      1 others
ORIGIN
Query Match      75.6%; Score 853; DB 13; Length 912;
Best Local Similarity 98.9%; Pred. No. 2.5e-211;
Matches 901; Conservative 0; Mismatches 5; Indels 5; Gaps 4;
QY 197 AGAACCCAGTCCCGGACCTGGG-CCTGGCAACATTAATTAATTCGCGGATCTGA 254
DB 1 AGAACCCAGTCCCGGACCTGGGCGCTGGGCGCAACATTAATTAATTCGCGGATCTGA 60
QY 255 TGGGGATCCCAAGCCCTGGTGCACCGTGTGGAAGAACCGGAGGCTGACGTGGAGTACTG 314

```

Db 61 TGGGATGCAAGCCCTGCTGTCACGCTGTAAGAACCGCAGGCTGAGCTGGAGTACTG 120  
Qy 315 TGATGTGCCCTCTGCTCCACCTGGCGCTGAGACAGTACAGCAGCCTCAGTTTCGAT 374  
Db 121 TGATGTGCCCTCTGCTCCACCTGGCGCTGAGACAGTACAGCAGCCTCAGTTTCGAT 180  
Qy 375 CAAAGAGGCTCTTTCGCGGACATCGCTCCACCCCTGGCAGGCTGCAATCTTTGCCAA 434  
Db 181 CAAAGAGGCTCTTTCGCGGACATCGCTCCACCCCTGGCAGGCTGCAATCTTTGCCAA 240  
Qy 435 GCACAGGAGTCCCGGAGAGCGCTCTGTCGGGGGGCATACATCAGTCTCTGCTG 494  
Db 241 GCACAGGAGTCCCGGAGAGCGCTCTGTCGGGGGGCATACATCAGTCTCTGCTG 300  
Qy 495 GATTCCTCTGCGCGCCACTGCTCCAGGAGAGGTTTCGCGCCCAACCACTGACGGTAT 554  
Db 301 GATTCCTCTGCGCGCCACTGCTCCAGGAGAGGTTTCGCGCCCAACCACTGACGGTAT 360  
Qy 555 CTTGGGAGAACATACCGGGTGGTCCCTGGCGAGAGGAGAGAAATTTGAAGTCGAAA 614  
Db 361 CTTGGGAGAACATACCGGGTGGTCCCTGGCGAGGAGGAGAGAAATTTGAAGTCGAAA 420  
Qy 615 ATACATTGTCCATAAGGAATTCGATGATGACATTAACGACATGACATTCGCTGCTGA 674  
Db 421 ATACATTGTCCATAAGGAATTCGATGATGACATTAACGACATGACATTCGCTGCTGA 480  
Qy 675 GCTGAATTCGGAATTCGTCGCGCTGTCGCGAGAGAGCGTGGTCCGCACTGTGTGCT 734  
Db 481 GCTGAATTCGGAATTCGTCGCGCTGTCGCGAGAGAGCGTGGTCCGCACTGTGTGCT 540  
Qy 735 TCCCCCGGAGACTGCGAGTCCCGACTGACCGAGTGTGAGTCTCCGCTACGGCAA 794  
Db 541 TCCCCCGGAGACTGCGAGTCCCGACTGACCGAGTGTGAGTCTCCGCTACGGCAA 600  
Qy 795 GCATGAGCCTTGTCTCTTCTTCTATTTCGAGCGGCTGAGAGGCTCATGTGAGTGTGA 854  
Db 601 GCATGAGCCTTGTCTCTTCTTCTATTTCGAGCGGCTGAGAGGCTCATGTGAGTGTGA 660  
Qy 855 CCCATCCAGCGCTGCACATCACAACTTTACTTAACAGAACAGTCAACCAACATGCT 914  
Db 661 CCCATCCAGCGCTGCACATCACAACTTTACTTAACAGAACAGTCAACCAACATGCT 720  
Qy 915 GTGTCTGGAGACATTCGAGCGGCGGCCCGCAAACTTTCACAGCAGCTGCGCAGG 974  
Db 721 GTGTCTGGAGACATTCGAGCGGCGGCCCGCAAACTTTCACAGCAGCTGCGCAGG 780  
Qy 975 -CGATTTCGGAGG-CCCCCTGGTGTCTGTAACGATGCGCGCATGCTTTGTTGGGCATC 1032  
Db 781 NCGATTTCGGAGGCCCCCTGGTGTCTGTAACGATGCGCGCATGCTTTGTTGGGCATC 840  
Qy 1033 ATCAGTGGGGCTGGGCTGTGGACAGAGGATGT-CCCCGGTGTGTACACAAAGTTTAC 1091  
Db 841 ATCAGTGGGGCTGGGCTGTGGACAGAGGATGTCCCCGGGTGTGTACACAAAGTTTAC 900  
Qy 1092 CAATACCTAG 1102  
Db 901 AAATACCTAG 911

RESULT 3  
LOCUS BX389609  
DEFINITION BX389609 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA  
clone CS0DI039YK15 5-PRIME, mRNA sequence.  
ACCESSION BX389609  
VERSION BX389609.1 GI:30463470  
KEYWORDS EST.  
SOURCE Homo sapiens  
ORGANISM Homo sapiens (human)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 962)

AUTHORS  
TITLE  
JOURNAL  
COMMENT

Li, W.B., Gruber, C., Jessee, J. and Polayes, D.  
Full-length cDNA libraries and normalization  
Unpublished  
Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 EVRY cedex - France  
Email: segre@genoscope.cns.fr, Web : www.genoscope.cns.fr  
was normalized. Library was constructed by Life Technologies, a  
division of Invitrogen. This sequence belongs to sequence cluster  
4498.r For more information about this cluster, see  
http://www.genoscope.cns.fr/  
cgi-bin/cluster.cgi?seq=CS0BAF025ZG09\_AF02376\_3&cluster=4498.r.  
Contact : Feng Liang Email : fliang@lifetech.com URL :  
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600  
Faraday Avenue Genoscope sequence ID : CS0BAF025ZG09\_AF02376\_3.

FEATURES  
source

1..962  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CS0DI039YK15"  
/tissue\_type="PLACENTA COT 25-NORMALIZED"  
/clone\_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"  
/note="1st strand cDNA was primed with a NotI-oligo(dT)  
primer. Five prime end enriched, double-strand cDNA was  
digested with Not I and cloned into the Not I and EcoR V  
sites of the pCMVSPORT 6 vector. Library was normalized."  
BASE COUNT 203 a 290 c 277 g 191 t 1 others  
ORIGIN

Query Match 75.4%; Score 850.2; DB 13; Length 962;  
Best Local Similarity 96.6%; Pred. No. 1.4e-210;  
Matches 879; Conservative 0; Mismatches 29; Indels 2; Gaps 1;

Qy 208 GCCCAGGACCTGGCGCTGGGCAACATTAATTAATTCGCGGATCTCTGATGGGGATGCCAAG 267  
Db 1 GCCCAGGACCTGGCGCTGGGCAACATTAATTAATTCGCGGATCTCTGATGGGGATGCCAAG 60  
Qy 268 CCCTGGTGCCACGTGCTGAAGAACCGCAGGCTGACGTGGGAGTACTGTGATGTGCCCTCC 327  
Db 61 CCCTGGTGCCACGTGCTGAAGAACCGCAGGCTGACGTGGGAGTACTGTGATGTGCCCTCC 120  
Qy 328 TGCTCCACCTGGCGCTTGAGACAGTACAGCAGCCTCAGTTTCGATCAAAAGAGGGCTC 387  
Db 121 TGCTCCACCTGGCGCTTGAGACAGTACAGCAGCCTCAGTTTCGATCAAAAGAGGGCTC 180  
Qy 388 TTGCGCGACATCGCCTCCACCCCTGGCAGGCTGCCATCTTTGCCAAGCAGAGGTCG 447  
Db 181 TTGCGCGACATCGCCTCCACCCCTGGCAGGCTGCCATCTTTGCCAAGCAGAGGTCG 240  
Qy 448 CCGGAGAGCGGTTCCCTGTCGGGGGCATCTCATCAGCTCCTGCTGATTTCTCTGCCC 507  
Db 241 CCGGAGAGCGGTTCCCTGTCGGGGGCATCTCATCAGCTCCTGCTGATTTCTCTGCCC 300  
Qy 508 GCCCAGCTTCCAGGAGAGGTTTCCGCCCCACCACTGACGCTGATCTTTGGGCGAACA 567  
Db 301 GCCCAGCTTCCAGGAGAGGTTTCCGCCCCACCACTGACGCTGATCTTTGGGCGAACA 360  
Qy 568 TACCGGCTGCTCCCTGGCGAGGAGCAGAAATTTGAAGTCGAAAATACATTTGTCAT 627  
Db 361 TACCGGCTGCTCCCTGGCGAGGAGCAGAAATTTGAAGTCGAAAATACATTTGTCAT 420  
Qy 628 AAGGAATTCGATGATGACCTTACGACAATGACATTCGCTGCTGACGCTGAAATCGAT 687  
Db 421 AAGGAATTCGATGATGACCTTACGACAATGACATTCGCTGCTGACGCTGAAATCGAT 480  
Qy 688 TCCTCCGCTGTGCCAGGAGAGCAGCTGGTCCGCACTGTGTGCTCTCCCGCGCGGAC 747  
Db 481 TCCTCCGCTGTGCCAGGAGAGCAGCTGGTCCGCACTGTGTGCTCTCCCGCGCGGAC 540  
Qy 748 CTGACGCTGCCGAGCTGGACGGAGTGTGAGCTCTCCGGCTACGGCAAGCATGAGGCTTG 807  
Db 541 CTGACGCTGCCGAGCTGGACGGAGTGTGAGCTCTCCGGCTACGGCAAGCATGAGGCTTG 600

QY 808 TCTCTTTCTATTTCGAGCGGCTGAAGAGGCTCATGTGACACTGTACCCATCCAGCGC 867  
 DB 601 TCTCTTTCTATTTCGAGCGGCTGAAGAGGCTCATGTGACACTGTACCCATCCAGCGC 660  
 QY 868 TGCACATCACAACTTTACTTTAAACAGAACAGTACCGACACATGCTGTGCTGGAGAC 927  
 DB 661 TGCACATCACAACTTTACTTTAAACAGAACAGTACCGACACATGCTGTGCTGGAGAC 720  
 QY 928 ACTCGAGCGGGGGCCCCAGGAACTTGACAGCGCTCGCAGGGCGATTCGGAGGC 987  
 DB 721 ACTCGAGCGGGGGCCCCAGGAACTTGACAGCGCTCGCAGGGCGATTCGGAGGC 780  
 QY 988 CCCCTGCTGTCTGAAGAGTGGCGCATGCTTTGTTGGCATCATCAGCTGGGGCTG 1047  
 DB 781 CCCCTGCTGTCTGAAGAGTGGCGCATGCTTTGTTGGCATCATCAGCTGGGGCTG 840  
 QY 1048 GGCTGTGACAGAGGATGTCCTGGGTGTGTACACAAAGGTTACCACTACCTAGACTGG 1107  
 DB 841 GGCTGTGACAGAGGATGTCCTGGGTGTGTACACAAAGGTTACCACTACCTAGACTGG 898  
 QY 1108 ATTCTGTACA 1117  
 DB 899 ATTCGGACA 908

RESULT 4  
 BUI46167  
 LOCUS  
 DEFINITION AGENCOURT\_7981262 Lupski\_dorsal\_root\_ganglion Homo sapiens cDNA  
 clone IMAGE:6185661 5', mRNA sequence.

ACCESSION BUI46167

VERSION BUI46167.1 GI:22659699

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 891)

NIH-MGC <http://mgc.nci.nih.gov/>.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished

Contact: Robert Strausberg, Ph.D.

Email: [cgapbs@mail.nih.gov](mailto:cgapbs@mail.nih.gov)

Tissue Procurement: Dr. James R. Lupski

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: LLNL3576 row: m column: 22

High quality sequence stop: 665.

Location/Qualifiers

1..891

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="caxon:9606"

/clone="IMAGE:6185661"

/sex="male"

/tissue type="dorsal root ganglia"

/dev stage="adult, 36 yr"

/lab host="DH10B"

/clone\_lib="Lupski dorsal root ganglion"

/note="Vector: pCMV-SPORT6 (Life Technologies); Site\_1:

Not; Site\_2: SalI; cDNA made by oligo-dr priming.

Directionally cloned using the following adaptors:

5'-TCGACCCACGGCTCCG-3' and

5'-GACTAGTCTAGATCGGAGCGGCCCT(15)-3'. Size selected >

1 kb for average insert length 1.7 kb. This is a primary

library, non-amplified. Library constructed by Life

Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor

College of Medicine) and is available through Life

BASE COUNT 178 a 266 c 271 g 176 t  
 ORIGIN

Query Match 72.6%; Score 818.6; DB 13; Length 891;  
 Best Local Similarity 98.2%; Pred. No. 2.3e-202;  
 Matches 849; Conservative 0; Mismatches 14; Indels 2; Gaps 2;

QY 202 CCACAGTCCCGCAGGACCTGGCGCCAAACATATTAATCTGCGGAATCTCTGATGGGAT 261  
 DB 1 CCCAGTCCCGCAGGACCTGGCGCCAAACATATTAATCTGCGGAATCTCTGATGGGAT 60  
 QY 262 GCCAAGCCCTGGTGCCACAGTGTGAAGAACCGCAGGCTGACGTGGAGTACTGTGATGTG 321  
 DB 61 GCCAAGCCCTGGTGCCACAGTGTGAAGAACCGCAGGCTGACGTGGAGTACTGTGATGTG 120  
 QY 322 CCCTCTCTCTCACTCCGCTGTGAGACAGTACAGCCAGCTCTGATTTGCGCATCAAGGA 381  
 DB 121 CCCTCTCTCTCACTCCGCTGTGAGACAGTACAGCCAGCTCTGATTTGCGCATCAAGGA 180  
 QY 382 GGGCTCTTCGCGCAGACATCGCTCCACCCCTGGCAGGCTGCCATCTTTGCCAAGCAGG 441  
 DB 181 GGGCTCTTCGCGCAGACATCGCTCCACCCCTGGCAGGCTGCCATCTTTGCCAAGCAGG 240  
 QY 442 AGGTGCGCCCGAGAGCGGTTCTGTGCGGGGCATCTCATCAGCTCTGTGATTTCTC 501  
 DB 241 AGGTGCGCCCGAGAGCGGTTCTGTGCGGGGCATCTCATCAGCTCTGTGATTTCTC 300  
 QY 502 TGTGCGCCCGCAGTCTTCCAGGAGAGGTTTCCGCCCCACACCTGACGGTATCTTGGGC 561  
 DB 301 TGTGCGCCCGCAGTCTTCCAGGAGAGGTTTCCGCCCCACACCTGACGGTATCTTGGGC 360  
 QY 562 AGAATACACCGGTGTGCTCCCTGGCAGGAGGAGAGAAATTTGAGTCGAAATATCATTT 621  
 DB 361 AGAATACACCGGTGTGCTCCCTGGCAGGAGGAGAGAAATTTGAGTCGAAATATCATTT 420  
 QY 622 GTCCATAAGGAATTCGATGATGACACTTACGACATGACATTTGGCTGTGACGTGAAA 681  
 DB 421 CTCCATAGGAATTCGATGATGACACTTACGACATGACATTTGGCTGTGACGTGAAA 480  
 QY 682 TCGGATTCGTCCTGCTGTCGCGAGGAGGAGGCTGTGTCGCACTGTGTCTTCCCGC 741  
 DB 481 TCGGATTCGTCCTGCTGTCGCGAGGAGGAGGCTGTGTCGCACTGTGTCTTCCCGC 540  
 QY 742 GCGGACCTGACGCTGCGGAGTGTGAGCTCTCCGCTACGCGCAAGCATGAG 801  
 DB 541 GCGGACCTGACGCTGCGGAGTGTGAGCTCTCCGCTACGCGCAAGCATGAG 600  
 QY 802 GCCTTGTCTCTTCTTATTCGAGCGGCTGAAAGAGGCTCATGTGACACTGTACCCATCC 861  
 DB 601 GCCTTGTCTCTTCTTATTCGAGCGGCTGAGAGGCTCATGTGACACTGTACCCATCC 660  
 QY 862 AGCGCTGACATCACAACTTTACTTTAAGAGAGTACCGGCAACATGCTGTGTGCT 921  
 DB 661 AGCGCTGACATCACAACTTTACTTTAAGAGAGTACCGGCAACATGCTGTGTGCT 720  
 QY 922 GGAGACACTCGGAGCGGCTGGGCCCCAGGCAAACTTGCAGACGCTTCCAGGGGATTC 980  
 DB 721 GGAGACACTCGGAGCGGCTGGGCCCCAGGCAAACTTGCAGACGCTTCCAGGGGATTC 780  
 QY 981 GGAGAGCGCCCTGTGTGTCTGAAACGATGGCCGCTCATGCTGTGGGATCATCAGCT 1039  
 DB 781 GGAGAGCGCCCTGTGTGTCTGAAACGATGGCCGCTCATGCTGTGGGATCATCAGCT 840  
 QY 1040 GGAGAGCGGCTGTGTGTGAGACAGGAAGA 1064  
 DB 841 GGAGAGCGCCCTGTGTGTGAGACAGGAAGA 865

RESULT 5

BQ278177

LOCUS

DEFINITION

BQ278177 993 bp mRNA linear EST 07-MAY-2002  
 AGENCOURT\_7061709 NIH\_MGC\_109 Homo sapiens cDNA clone IMAGE:5805153

```

5', mRNA sequence.
ACCESSION BQ278177 1 GI:20488385
VERSION BQ278177.1
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
JOURNAL NIH-MGC http://mgi.nci.nih.gov/.
COMMENT National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-re@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM2043 row: 0 column: 10
High quality sequence stop: 681.
FEATURES
source
Location/Qualifiers
1..993
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5805153"
/tissue_type="teratocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_109"
/Note="Organ: ovary; Vector: pOTB7; Site 1: EcoRI; Site 2:
XhoI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC Library."
BASE COUNT 235 a 295 c 269 g 194 t
ORIGIN

Query Match 72.1%; Score 813.8; DB 13; Length 993;
Best Local Similarity 99.6%; Pred. No. 4.3e-201;
Matches 826; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

Qy 300 GACGTGGAGTACTGTGATGTCCTCTCTGCTCCACCTGGCGCTGACAGTACAGCCA 359
Db 1 GACGTGGAGTACTGTGATGTCCTCTCTGCTCCACCTGGCGCTGACAGTACAGCCA 59

Qy 360 GCCTCAGTTTCGATCAAGAGGAGCTCTTCGCGACATCGCTCCACCTGGCAGGC 419
Db 60 GCCTCAGTTTCGATCAAGAGGAGCTCTTCGCGACATCGCTCCACCTGGCAGGC 119

Qy 420 TGCCATCTTTCCAGCAGAGAGTCCCGGAGAGCGGTTCCTGTGCGGGGGCTACT 479
Db 120 TGCCATCTTTCCAGCAGAGAGTCCCGGAGAGCGGTTCCTGTGCGGGGGCTACT 179

Qy 480 CATCAGCTCTCTGATTTCTCTGCGCGCCACTGTTTCCAGAGAGGTTTCGCGCCA 539
Db 180 CATCAGCTCTCTGATTTCTCTGCGCGCCACTGTTTCCAGAGAGGTTTCGCGCCA 239

Qy 540 CCACCTGACGTGATCTTGGGAGAGATACCGGTGCTGCTCCGAGGAGGAGGAGCA 599
Db 240 CCACCTGACGTGATCTTGGGAGAGATACCGGTGCTGCTCCGAGGAGGAGGAGCA 299

Qy 600 ATTTGAAGTCAAAAATACATTGTTCATTAAGAAATTCGATGATGACACTTACGACAATGA 659
Db 300 ATTTGAAGTCAAAAATACATTGTTCATTAAGAAATTCGATGATGACACTTACGACAATGA 359

Qy 660 CATTGCGCTGCTGAGCTGAATCGGATTCCTCGCTGTGCCAGGAGACAGCGTGGT 719
Db 1 CATTGCGCTGCTGAGCTGAATCGGATTCCTCGCTGTGCCAGGAGACAGCGTGGT 419

Qy 720 CCGCACTGTGTGCTTCCCGCGGAGCTGCGAGCTGCGAGCTGCGAGCTGAGCT 779
Db 420 CCGCACTGTGTGCTTCCCGCGGAGCTGCGAGCTGCGAGCTGCGAGCTGAGCT 479

Qy 780 CTCCGGCTACGGCAAGCATGAGGCTTGTCTCTTTTATTTCGAGCGGCTGAAGAGGC 839
Db 480 CTCCGGCTACGGCAAGCATGAGGCTTGTCTCTTTTATTTCGAGCGGCTGAAGAGGC 539

Qy 840 TCATGTGAGCTGACCCATCCAGCCCTGCACATCAACATTTACTTACAGAACAGT 899
Db 540 TCATGTGAGCTGACCCATCCAGCCCTGCACATCAACATTTACTTAAACAGAACAGT 599

Qy 900 CACGCAACATGCTGTGTCTGAGACACTTCGAGCGGCGGCGCCAGCAAACTTCA 959
Db 600 CACGCAACATGCTGTGTCTGAGACACTTCGAGCGGCGGCGCCAGCAAACTTCA 659

Qy 960 CGAGCCTGCGAGGCGGATTCGAGGAGCCCTCTGCTGTCTGAACGATGCCGATGAC 1019
Db 660 CGAGCCTGCGAGGCGGATTCGAGGAGCCCTCTGCTGTCTGAACGATGCCGATGAC 719

Qy 1020 TTTGCTGGCATCATCAGCTGGGCGCTGGGAGAGGATGTCCCGGCTGTGTA 1079
Db 720 TTTGCTGGCATCATCAGCTGGGCGCTGGGAGAGGATGTCCCGGCTGTGTA 779

Qy 1080 CACAAAGTTACCAACTACCTAGACTGATTCGTGACAAACATGCGACCG 1128
Db 780 CACAAAGTTACCAACTACCTAGACTGATTCGTGACAAACATGCGACCG 828

RESULT 6
BO690978 883 bp mRNA linear EST 15-JUL-2002
LOCUS AGENCOURT_8046707 NIH_MGC_110 Homo sapiens cDNA clone IMAGE:6209120
DEFINITION 5', mRNA sequence.
ACCESSION BO690978
VERSION BO690978.1 GI:21816294
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-re@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM2367 row: 0 column: 09
High quality sequence stop: 558.
FEATURES
source
Location/Qualifiers
1..883
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6209120"
/tissue_type="ductal carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_110"
/Note="Organ: pancreas; Vector: pOTB7; Site 1: XhoI;
Site 2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCAGCAG(G). Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit

```

```
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."
BASE COUNT      184 a      265 c      259 g      175 t
ORIGIN
Query Match      71.9%; Score 811.2; DB 13; Length 883;
Best Local Similarity 99.0%; Pred. No. 1.9e-200;
Matches 816; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
Qy 178 GCGAAGTTTACACAGCAGACAGACCCAGTGGCCAGGCACTGGGCTGGGCAACATAT 237
Db 1 GCGAAGTTTACACAGCAGACAGACCCAGTGGCCAGGCACTGGGCTGGGCAACATAT 60
Qy 238 TACTGCCGGAATCCTGATGGGATGCCAAGCCCTGGTCCACAGTCTCTGAAGAACCGCAGG 297
Db 61 TACTGCCGGAATCCTGATGGGATGCCAAGCCCTGGTCCACAGTCTCTGAAGAACCGCAGG 120
Qy 298 CTGACGTGGGAGTACTGTGATGTGCCCTCTCTGCTCCACCTGGGCTGAGACAGTACAGC 357
Db 121 CTGACGTGGGAGTACTGTGATGTGCCCTCTCTGCTCCACCTGGGCTGAGACAGTACAGC 180
Qy 358 CAGCCTCAGTTTCGATCAAAAGGAGGCTCTTCGCCGACATCGCTCCACCCCTGGCAG 417
Db 181 CAGCCTCAGTTTCGATCAAAAGGAGGCTCTTCGCCGACATCGCTCCACCCCTGGCAG 240
Qy 418 GCTGCCATCTTTGCCAAGCACAGGAGGTGCGCCCGGAGAGCGGTTCTGTGGGGGCATA 477
Db 241 GCTGCCATCTTTGCCAAGCACAGGAGGTGCGCCCGGAGAGCGGTTCTGTGGGGGCATA 300
Qy 478 CTATPACGTCTGTGATGTTCTCTCTCGGCCCATCTGCTTCCAGGAGAGTTTCCGCC 537
Db 301 CTATPACGTCTGTGATGTTCTCTCTCGGCCCATCTGCTTCCAGGAGAGTTTCCGCC 360
Qy 538 CACCACTTCAGGTGATCTTGGGAGAACATACCGGGTGGTCCCTGGCGAGGAGGAGC 597
Db 361 CACCACTTCAGGTGATCTTGGGAGAACATACCGGGTGGTCCCTGGCGAGGAGGAGC 420
Qy 598 AAATTTGAAGTGGAAATACATTTGTCATAGGAATTCGATGATGACACTTACGACAA 657
Db 421 AAATTTGAAGTGGAAATACATTTGTCATAGGAATTCGATGATGACACTTACGACAA 480
Qy 658 GACATTTGGCTGCTGACGTGAATTCGGATTCGTCCCGCTGTGCCAGGAGAGCAGCGTG 717
Db 481 GACATTTGGCTGCTGACGTGAATTCGGATTCGTCCCGCTGTGCCAGGAGAGCAGCGTG 540
Qy 718 GTCCGCACTGTGCTCTTCCCGCGGACCTGACGTGCGGACTGACGAGGTGAG 777
Db 541 GTCCGCACTGTGCTCTTCCCGCGGACCTGACGTGCGGACTGACGAGGTGAG 600
Qy 778 CTCTCCGCTACGGCAAGCATGAGGCTTGTCTCTCTTCTATTCGGAGCGCTGAAGGAG 837
Db 601 CTCTCCGCTACGGCAAGCATGAGGCTTGTCTCTCTTCTATTCGGAGCGCTGAAGGAG 660
Qy 838 GCTCATGTGACAGTGTACCCATCCAGCGCTGCAATCAACAATTTACTTAACAGAAC 897
Db 661 GCTCATGTGACAGTGTACCCATCCAGCGCTGCAATCAACAATTTACTTAACAGAAC 720
Qy 898 GTACCCGACAACTGCTGTGCTGTGAGACACTCGAGGCGGCGGCCCAACTTG 957
Db 721 GTACCCGACAACTGCTGTGCTGTGAGACACTCGAGGCGGCGGCCCAACTTG 780
Qy 958 CACGAGCGCTGCGAGGGGATTCGGGAGGCGCCCTGTGTGCT 1001
Db 781 CACGAGCGCTGCGAGGGGATTCGGGAGGCGCCCTGTGTGCT 824

RESULT 7
LOCUS      BX389608
DEFINITION BX389608 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
            clone CS0D1039YK15 5-PRIME, mRNA sequence.
ACCESSION  BX389608
VERSION    BX389608.1 GI:30463469
```

```
KEYWORDS EST. Homo sapiens (human)
SOURCE     Homo sapiens
ORGANISM   Homo sapiens
REFERENCE  1 (bases 1 to 918)
AUTHORS    Li,W.B., Gruber,C., Jesses,J. and Polayes,D.
TITLE      Full-length cDNA libraries and normalization
JOURNAL    Unpublished
COMMENT     Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr; Web : www.genoscope.cns.fr
            was normalized. Library was constructed by Life Technologies, a
            division of Invitrogen. This sequence belongs to sequence cluster
            4498.r For more information about this cluster, see
            http://www.genoscope.cns.fr/
            cgi-bin/cluster.cgi?seq=CS0BAP025ZG09_AF02376_2&cluster=4498.r.
            Contact : Feng Liang Email : fliang@lifetech.com URL :
            http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
            Paraday Avenue Genoscope sequence ID : CS0BAP025ZG09_AF02376_2.
FEATURES   Location/Qualifiers
            source          1..918
                        /organism="Homo sapiens"
                        /mol_type="mRNA"
                        /db_xref="taxon:9606"
                        /clone="CS0D1039YK15"
                        /tissue_type="PLACENTA COT 25-NORMALIZED"
                        /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
                        /note="1st strand cDNA was primed with a NotI-oligo (dr)
                        primer. Five prime end enriched, double-strand cDNA was
                        digested with Not I and cloned into the Not I and Ecor V
                        sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT      189 a      273 c      264 g      186 t
ORIGIN
Query Match      71.6%; Score 807.6; DB 13; Length 918;
Best Local Similarity 96.4%; Pred. No. 1.7e-199;
Matches 855; Conservative 0; Mismatches 29; Indels 3; Gaps 3;
Qy 236 ATTACTCCGGAATCTCTGATGGGATGCCAAGCCCTGGTGCACGTGCTGAAGAACCGCA 295
Db 17 ATTACTCCGGAATCTCTGATGGGATGCCAAGCCCTGGTGCACGTGCTGAAGAACCGCA 76
Qy 296 GGCTGACGTGGGAGTACTGTGATGTGCCCTCTCTGCTCCACTCGGCGCTGAGACAGTACA 355
Db 77 GGCTGACGTGGGAGTACTGTGATGTGCCCTCTCTGCTCCACTCGGCGCTGAGACAGTACA 136
Qy 356 GCCAGCCTCAGTTTCGATCAAAAGGAGGGTCTTTCGCCGACATCGCCTCCACCCCTGGC 415
Db 137 GCCAGCCTCAGTTTCGATCAAAAGGAGGGTCTTTCGCCGACATCGCCTCCACCCCTGGC 196
Qy 416 AGGCTGCCATCTTTGCCAAGCACAGGAGGTGCGCCGAGAGCGGTTCCTGTGCGGGGCA 475
Db 197 AGGCTGCCATCTTTGCCAAGCACAGGAGGTGCGCCGAGAGCGGTTCCTGTGCGGGGCA 256
Qy 476 TACTCATCAGTCTCTGTGATTTCTCTGCGCCCACTGCTTCCAGGAGAGTTTCGCG 535
Db 257 TACTCATCAGTCTCTGTGATTTCTCTGCGCCCACTGCTTCCAGGAGAGTTTCGCG 316
Qy 536 CCCACCACTTCAGCGGTGATCTTGGGCGAACAATACCGGGTGGTCCCTGGCGAGGAGGAGC 595
Db 317 CCCACCACTTCAGCGGTGATCTTGGGCGAACAATACCGGGTGGTCCCTGGCGAGGAGGAGC 376
Qy 596 AGAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA 655
Db 377 AGAAATTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA 436
Qy 656 ATGACATTGGCTGCTCAGCTGAATTCGGATTTCGTCCTGCTGTCGCCAGGAGAGCAGCG 715
Db 437 ATGACATTGGCTGCTCAGCTGAATTCGGATTTCGTCCTGCTGTCGCCAGGAGAGCAGCG 496
Qy 716 TGTGCCGCACTGTGTGCTTTCCCGCGGACCTGACAGCTGCCGAGTGGACGAGGTGTG 775
```





Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

NIH-MGC <http://mgc.nci.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished  
Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)

Tissue Procurement: ATCC  
CDNA Library Preparation: Rubin Laboratory

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>  
Plate: LLC2550 row: e column: 21

High quality sequence stop: 630.  
Location/Qualifiers

1..888  
/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:6084476"

/tissue\_type="ductal carcinoma, cell line"

/lab\_host="DH10B (phage-resistant)"

/clone\_lib="NIH MGC 110"

/note="Organ: pancreas; Vector: pOTB7; Site 1: XhoI;  
Site 2: EcoRI; cDNA made by oligo-dT priming.

Directionally cloned into EcoRI/XhoI sites using the  
following 5' adaptor: GGCACGAG(G). Library constructed by

Ling Hong in the laboratory of Gerald M. Rubin (University  
of California, Berkeley) using ZAP-cDNA synthesis kit

(Stratagene, Berkeley) and Superscript II RT (Life Technologies).  
Note: this is a NIH\_MGC Library."

BASE COUNT 181 a 265 c 270 g 172 t

Query Match 69.4%; Score 782.8; DB 13; Length 888;  
Best Local Similarity 97.2%; Pred. No. 4,9e-193;

Matches 861; Conservative 0; Mismatches 17; Indels 8; Gaps 6;

QY 178 GCGAGGTTTACACAGCAGAACCCAGTGCCTCCAGGCACTGGGCTGGGCAACATAAT 237

Db 1 GCGAGGTTTACACAGCAGAACCCAGTGCCTCCAGGCACTGGGCTGGGCAACATAAT 60

QY 238 TACTGCCGATCTGTATGGGATGCCAAGCCCTGCTCCAGGCTGCTGAAGACCGCAGG 297

Db 61 TACTGCCGATCTGTATGGGATGCCAAGCCCTGCTCCAGGCTGCTGAAGACCGCAGG 120

QY 298 CTGAGTGGGAGTACTGTGATGTGCTCTCTCCAGTCCAGGCTGAGACAGTACAGC 357

Db 121 CTGAGTGGGAGTACTGTGATGTGCTCTCTCCAGTCCAGGCTGAGACAGTACAGC 180

QY 358 CAGCTCAGTTTGCATCAAGAGGGGCTTTGGCCGACATCGCTCCACCCCTGGCAG 417

Db 181 CAGCTCAGTTTGCATCAAGAGGGGCTTTGGCCGACATCGCTCCACCCCTGGCAG 240

QY 418 GCTGCATCTTTGCCAAGCAGAGGCTGCGCCGAGAGCGGTTCTGTGCGGGGCATA 477

Db 241 GCTGCATCTTTGCCAAGCAGAGGCTGCGCCGAGAGCGGTTCTGTGCGGGGCATA 300

QY 478 CTCATCAGTCTCTGTGGATTCTCTCTCCGCGCCACTCTCTCCAGGAGAGGTTTCCGCC 537

Db 301 CTCATCAGTCTCTGTGGATTCTCTCTCCGCGCCACTCTCTCCAGGAGAGGTTTCCGCC 360

QY 538 CACCACTGACGGTGTATTTGGCGAGAACATACCGGGTGGTCCCTGGCGAGGAGGAGCAG 597

Db 361 CACCACTGACGGTGTATTTGGCGAGAACATACCGGGTGGTCCCTGGCGAGGAGGAGCAG 420

QY 598 AAATTTGAATCGAAATATATGTTCCATAGGATTCGATGATGACACTTACGACAT 657

Db 421 AAATTTGAATCGAAATATATGTTCCATAGGATTCGATGATGACACTTACGACAT 480

QY 658 GACATTCGCTGTGCTGACAGCTGAATATCGATTCGCTCCGCTGTGCTCCAGGAGCAGCGTG 717

Db 481 GACATTCGCTGTGCTGACAGCTGAATATCGATTCGCTCCGCTGTGCTCCAGGAGCAGCGTG 540

QY 718 GTCCGCACTGTGTGCTTCCCGCGGACCTGCGAGTCCGCGAGTCCGAGTCCGAGTGTGAG 777

Db 541 GTCCGCACTGTGTGCTTCCCGCGGACCTGCGAGTCCGCGAGTCCGAGTGTGAG 600

QY 778 CTCTCCGGCTACCGCAGCATGAGGCTTCTCTCTTTCTATTCTGGAGCGCTGAAGGAG 837

Db 601 CTCTCCGGCTACCGCAGCATGAGGCTTCTCTCTTTCTATTCTGGAGCGCTGAAGGAG 660

QY 838 GCTCATGTGAGACTGTACCCATCCAGCGCTGCACATCAACAATTTACTTTAAACAGAAC 897

Db 661 GCTCATGTGAGACTGTACCCATCCAGCGCTGCACATCAACAATTTACTTTAAACAGAAC 720

QY 898 GTACCCGACAACTGCTGTGTGCTGGAGACACTCGG--AGCGCGGGGCCCCCAGGCAAC 954

Db 721 GTACCCGACAACTGCTGTGTGCTGGAGACACTCGG--AGCGCGGGGCCCCCAGGCAAC 780

QY 955 TTGCACGACGCGCT--GCCAGGGCGATTTCGGAGGCCCCCTCGT-GTGTCTGAACGATGG-C 1011

Db 781 TTGCACGACGCGCT--GCCAGGGCGATTTCGGAGGCCCCCTCGT-GTGTCTGAACGATGG-C 840

QY 1012 CGCATGACTTT--GGTGGGCGATCATCAGCTGGGGCTGGGGCTGTGG 1055

Db 841 CGCATGACTTTGGGGGGGATCATCAACTGGGGCTGGGGCTGTGG 886

RESULT 10  
BQ927899

LOCUS BQ927899 950 bp mRNA linear EST 20-AUG-2002

DEFINITION AGENCOURT 8775988 NIH\_MGC\_18 Homo sapiens cDNA clone IMAGE:6373489

5', mRNA sequence.

ACCESSION BQ927899

VERSION BQ927899.1 GI:22342930

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

NIH-MGC <http://mgc.nci.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished  
Contact: Robert Strausberg, Ph.D.

Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
Tissue Procurement: DCTD/DTF/Gazdar

CDNA Library Preparation: Rubin Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>

Plate: LLC2550 row: p column: 02  
High quality sequence start: 35

High quality sequence stop: 684.  
Location/Qualifiers

1..950  
/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:6373489"

/tissue\_type="large cell carcinoma"

/lab\_host="DH10B (phage-resistant)"

/clone\_lib="NIH MGC 18"

/note="Organ: lung; Vector: pOTB7; Site 1: XhoI; Site 2:  
EcoRI; cDNA made by oligo-dT priming. Directionally cloned  
into EcoRI/XhoI sites using the following 5' adaptor:  
GGCAGAG(G). Library constructed by Ling Hong in the  
laboratory of Gerald M. Rubin (University of California,  
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and  
Superscript II RT (Life Technologies). Note: this is a

FEATURES

source

BASE COUNT	207 a	283 c	279 g	181 t
ORIGIN	NIH_MGC Library."			
Query Match	67.8%; Score 765.2; DB 13; Length 950;			
Best Local Similarity	98.1%; Pred. No. 2e-188;			
Matches	806; Conservative 0; Mismatches 13; Indels 3; Gaps 3;			
Qy	276	CCACGTGCTGAAGAACCGCAGCGTGAAGTGGAGTACTGTGATGTGCGCTCTCTGTCTCCAC	335	
Db	43	CCCCGTGCTGAAGAACCGCAGCGTGAAGTGGAGTACTGTGATGTGCGCTCTCTGTCTCCAC	102	
Qy	336	CTGCGGCTGAGACAGTACAGCAGCCTCAGTTTCGATCAAGAGAGGCTCTTCGCGGA	395	
Db	103	CTGCGGCTGAGACAGTACAGCAGCCTCAGTTTCGATCAAGAGAGGCTCTTCGCGGA	162	
Qy	396	CATCGCTCCCAACCCCTGGCAGCTGCCATCTTTGCCAAGCACAGGAGTGCCTCGGAGA	455	
Db	163	CATCGCTCCCAACCCCTGGCAGCTGCCATCTTTGCCAAGCACAGGAGTGCCTCGGAGA	222	
Qy	456	CGGTTCTGTGTCGGGGGCATCTCATCAGCTCTCTGTGATTTCTCTCTGCGGCCACTG	515	
Db	223	CGGTTCTGTGTCGGGGGCATCTCATCAGCTCTCTGTGATTTCTCTCTGCGGCCACTG	282	
Qy	516	CTTCAGAGAGG-TTTCGGCCCAACCACTGACGTGATCTTGGGAGAACATACCGGG	574	
Db	283	CTTCAGAGAGGTTTCTGCCCCACCACTGACGTGATCTTGGGAGAACATACCGGG	342	
Qy	575	TGGTCCCTGGCAGAGAGAGAGAAATTTGAAGTCGAAATATACATTCTCCATAGGAAT	634	
Db	343	TGGTCCCTGGCAGAGAGAGAGAAATTTGAAGTCGAAATATACATTCTCCATAGGAAT	402	
Qy	635	TCGATGATGACATTTACGACATGACATTTGGCTGCTGACGTGAAATTCGATTCGTC	694	
Db	403	TCGATGATGACATTTACGACATGACATTTGGCTGCTGACGTGAAATTCGATTCGTC	462	
Qy	695	GCTGTGCCAGAGAGAGCGTGTGTCGACATGTGTGCTTCCCGCGGACCTGCAGC	754	
Db	463	GCTGTGCCAGAGAGAGCGTGTGTCGACATGTGTGCTTCCCGCGGACCTGCAGC	522	
Qy	755	TGCCGACTGACGAGGTGACGTCTCCGCTACGGCAAGCATGAGCCTTGTCTCTT	814	
Db	523	TGCCGACTGACGAGGTGACGTCTCCGCTACGGCAAGCATGAGCCTTGTCTCTCTT	582	
Qy	815	TCTATTGCGAGCGGTGAAGAGGCTCATGTGACAGCTGTACCCATCCAGCGCTGCACAT	874	
Db	583	TCTATTGCGAGCGGTGAAGAGGCTCATGTGACAGCTGTACCCATCCAGCGCTGCACAT	642	
Qy	875	CACAACTTTACTTAAACAGAAAGTACCGACAGTCTGTGTGCTGGAGACACTCGGA	934	
Db	643	CACAACTTTACTTAAACAGAAAGTACCGACAGTCTGTGTGCTGGAGACACTCGGA	702	
Qy	935	CGCGGGCCCCAGCAAACTTGCACAGCGCTCCAGGGGATTCGGAGGCCCCCTGG	994	
Db	703	CGCGGGCCCCAGCAAACTTGCACAGCGCTCCAGGGGATTCGGAGGCCCCCTGG	762	
Qy	995	TGTGTCTGAACGATGGCCGCTGATCTTTGTGGGCTCATCAG-CTGGGGCTTGGGCTGT	1053	
Db	763	TGTGTCTGAACGATGGCCGCTGATCTTTGTGGGCTCATCAGCTTGGGGCTTGGGCTGT	822	
Qy	1054	GGACAG-AAGGATGTCCGGGTGTACACAAAGGTTACCAA	1094	
Db	823	GGACAGAAAGGCCCCGGGTGGGACACCAAGGTTACCCA	864	
RESULT 11				
BI765247				
LOCUS	603050338F1 NIH_MGC_116 Homo sapiens cDNA clone IMAGE:5190656 5',			
DEFINITION	mRNA sequence.			
ACCESSION	BI765247			
VERSION	BI765247.1 GI:15756825			
KEYWORDS	EST.			

Source

1. 916

Location/Qualifiers

organism="Homo sapiens"

mol\_type="mRNA"

db\_xref="taxon:9606"

clone="IMAGE:5190656"

lab\_host="DH10B"

clone\_lib="NIH MGC 116"

notes="Organ: pooled colon, kidney, stomach; Vector: PCMV-SPORT6; Site 1: NotI; Site 2: SmaI; Source anonymous pool of 3 colons, age 26 yo male, 49 yo female, 71 yo male and 70 yo female. Library is stomachs, 62 yo male and directionally cloned (EcoRV site is oligo-dr primed and destroyed upon cloning). Average insert size 1.4 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 023. Note: this is a NIH MGC Library."

Base Count

204 a 272 c 256 g 184 t

Query Match

67.8%; Score 764.4; DB 12; Length 916;

Best Local Similarity

99.7%; Pred. No. 3.1e-188;

Matches

776; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

Qy

352

TACAGCAGCCTCAGTTTCGCATCAAGAGGCTCTTCGCGACATCGCTCCACCCC

411

|||||

Db

1

TACAGCAGCCTCAGTTTCGCATCAAGAGGCTCTTCGCGACATCGCTCCACCCC

60

|||||

Qy

412

TGGCAGGCTGCCATCTTTGCCAAGCACAGGAGTTCGCCGAGAGCGGTTCTGTGCGG

471

|||||

Db

61

TGGCAGGCTGCCATCTTTGCCAAGCACAGGAGTTCGCCGAGAGCGGTTCTGTGCGG

120

|||||

Qy

472

GGCATCTCATAGCTCTCTGTGATTTCTCTGCGCGCCACATGCTTCCAGAGAGGTTT

531

|||||

Db

121

GGCATCTCATAGCTCTCTGTGATTTCTCTGCGCGCCACATGCTTCCAGAGAGGTTT

180

|||||

Qy

532

CCGCCCCACCATGAGGCTTGGGAGAGATACACCGGCTGCTCCGCGAGGAG

591

|||||

Db

181

CCGCCCCACCATGAGGCTTGGGAGAGATACACCGGCTGCTCCGCGAGGAG

240

|||||

Qy

592

GAGCAGAAATTTGAAGTCGAAATATATTTGTCATTAAGGAATTCGATGATGACATTAC

651

|||||

Db

241

GAGCAGAAATTTGAAGTCGAAATATATTTGTCATTAAGGAATTCGATGATGACATTAC

300

|||||

Qy

652

GACAAATGACATTCGCTGCTGAGTGAATTCGATTCCTGCTGCTCCAGGAGC

711

|||||

Db

301

GACAAATGACATTCGCTGCTGAGTGAATTCGATTCCTGCTGCTCCAGGAGC

360

|||||

Qy

712

AGCGTGTGCGCACTGTGCTGCTTCCCGCGGAGCTGCGGCTGCGGAGGAG

771

|||||

Db

361

AGCGTGTGCGCACTGTGCTGCTTCCCGCGGAGCTGCGGAGGAGCTGCGGAGGAG

420

|||||

QY 772 TGTGAGCTCTCGGTACGCAAGCATGAGCGCTTGTCTCTTTCTATTTCGAGCGGTG 831  
 Db 421 TGTGAGCTCTCGGTACGCAAGCATGAGCGCTTGTCTCTTTCTATTTCGAGCGGTG 480  
 QY 832 AAGGAGGCTCATGTGAGCTGTACCCATCCAGCGCTGCAATCAACAATTTACTTAAAC 891  
 Db 481 AAGGAGGCTCATGTGAGCTGTACCCATCCAGCGCTGCAATCAACAATTTACTTAAAC 540  
 QY 892 AGAACAGTCAACGCAACATGCTGTGTGTGAGACATCCGAGCGCGGCGCCAGGCA 951  
 Db 541 AGAACAGTCAACGCAACATGCTGTGTGTGAGACATCCGAGCGCGGCGCCAGGCA 600  
 QY 952 AACTTGCACGAGCGCTGCCAGGCGGATTCGGAGGCGCCCTGGTGTCTGAAACGATGSC 1011  
 Db 601 AACTTGCACGAGCGCTGCCAGGCGGATTCGGAGGCGCCCTGGTGTCTGAAACGATGSC 660  
 QY 1012 GCATGACTTTGGTGGGCAATCATAGCTGGGGCTTGGGCTGTGGACAGAGATGTCCCG 1071  
 Db 661 GCATGACTTTGGTGGGCAATCATAGCTGGGGCTTGGGCTGTGGACAGAGATGTCCCG 720  
 QY 1072 GGTGTGTACAAAGGTTACCAACTACCTAGAC-TGGATTGTCGACACATGCGACCG 1128  
 Db 721 GGTGTGTACAAAGGTTACCAACTACCTAGACTTGGATTGTCGACACATGCGACCG 778

RESULT 12  
 BE616613  
 LOCUS 601278864F1 NIH\_MGC\_39 Homo sapiens cDNA clone IMAGE:3611007 5',  
 DEFINITION mRNA sequence.

ACCESSION BE616613  
 VERSION BE616613.1 GI:9898212

KEYWORDS EST.  
 SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 948)  
 AUTHORS NIH-MGC <http://mgi.nci.nih.gov/>  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished

COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: [cgabbs-r@mail.nih.gov](mailto:cgabbs-r@mail.nih.gov)  
 Tissue Procurement: ATCC

cDNA Library Preparation: Ling Hong/Rubin Laboratory  
 DNA Sequencing Arrayed by: The I.M.A.G.E. Consortium (LLNL)

Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>  
 Plate: LCM267 row: p column: 16  
 High quality sequence start: 2  
 High quality sequence stop: 759.

Location/Qualifiers  
 1..948

/organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:3611007"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="NIH MGC 39"  
 /note="Organ: pancreas; Vector: pOTB7; Site 1: XhoI;  
 Site 2: EcoRI; cDNA made by oligo-dT priming.  
 Directionally cloned into EcoRI/XhoI sites using the  
 following 5' adaptor: GGCAGGAG(G). Library constructed  
 by Ling Hong in the laboratory of Gerald M. Rubin  
 (University of California, Berkeley) using ZAP-cDNA  
 synthesis kit (Stratagene) and Superscript II RT (Life  
 Technologies)."

173 a 295 c 293 g 187 t

BASE COUNT  
 ORIGIN

FEATURES

source

Query Match 62.7%; Score 706.8; DB 10; Length 948;  
 Best Local Similarity 93.3%; Pred. No. 3.2e-173;  
 Matches 793; Conservative 0; Mismatches 52; Indels 5; Gaps 5;  
 QY 72 GGGAAACAGTGAATGCTTACT-TTGGGAATGGGTGAGCTACCTACCTGGGACGCGACAGCTCA 130  
 Db 8 GAGGAACAGTGAATGCTTACTTATTTGGGAATGGGTGAGCTACCTGGGACGCGACAGCTCA 67  
 QY 131 CCGAGTCCGGTGGCTCTGCTGCTCC-CTGGAAATCCATGATCTCTGATAGCAAGGTTTAC 189  
 Db 68 CCGAGTCCGGTGGCTCTGCTGCTCCCGGTGGAAATCCATGATCTCTGATAGCAAGGTTTAC 127  
 QY 190 ACAGCACAGAACCCCACTGTCAGGCACTGGGCTGGGCAAAACATAATATTACTGCGGGAAT 249  
 Db 128 ACAGCACAGAACCCCACTGTCAGGCACTGGGCTGGGCAAAACATAATATTACTGCGGGAAT 187  
 QY 250 CTTGATGGGATGTCACAAAGCCCTGGTGCACGCTGTCGAAGAACCCGACGGCTGACCTGGGAG 309  
 Db 188 CTTGATGGGATGTCACAAAGCCCTGGTGCACGCTGTCGAAGAACCCGACGGCTGACCTGGGAG 247  
 QY 310 TACTGTGATGTCCTCTGCTCCACTGGGCTGGAGCAGTACACGCGACGCTCAGTTT 369  
 Db 248 TACTGTGATGTCCTCTGCTCCACTGGGCTGGAGCAGTACACGCGACGCTCAGTTT 307  
 QY 370 CGCATCAAGAGAGGGCTCTTTGCGCCGACATCGCTCCCACTGGGACGGCTGCCATCTTT 429  
 Db 308 CGCATCAAGAGAGGGCTCTTTGCGCCGACATCGCTCCCACTGGGACGGCTGCCATCTTT 367  
 QY 430 GCCAAGCACAGGAGGTGCGCCGAGAGCGGTTCTGTGCGGGGGCATACTCATCAGCTCC 489  
 Db 368 GCCAAGCACAGGAGGTGCGCCGAGAGCGGTTCTGTGCGGGGGCATACTCATCAGCTCC 427  
 QY 490 TGCTGGATTCTCTCTGCGCCGCTGCTCCAGGAGGTTTCGGCCGACCACTGACG 549  
 Db 428 TGCTGGATTCTCTCTGCGCCGCTGCTCCAGGAGGTTTCGGCCGACCACTGACG 487  
 QY 550 GTGATCTTTGGGCAAGAACATACCGGGTGGTCCCTGGGAGGAGGAGCAAGAAATTTGAAGTC 609  
 Db 488 GTGATCTTTGGGCAAGAACATACCGGGTGGTCCCTGGGAGGAGGAGCAAGAAATTTGAAGTC 547  
 QY 610 GAAAAATACATTTGTCATAAGGAATTCGATGATGACACTTACGACAAATGACATTTGCCGTG 669  
 Db 548 GAAAAATACATTTGTCATAAGGAATTCGATGATGACACTTACGACAAATGACATTTGCCGTG 607  
 QY 670 CTCGAGCTGAAATCGGATTCGTCCTGCTGCGCCGAGGAGGAGCGGTGGTCCGCACTGTG 729  
 Db 608 CTCGAGCTGAAATCGGATTCGTCCTGCTGCGCCGAGGAGGAGCGGTGGTCCGCACTGTG 667  
 QY 730 TGCCTTCCCGCGGACCTGCGAGCTGCGGAGTGGAGCGAGTGTGAGCTCTCCGGCTAC 789  
 Db 668 TGCCTT-CCCGCGGAGCTGCGAGTGGAGCGGAGGTTGAGCTCTCCGGCTAC 726  
 QY 790 GGCAAGCATGAGGCTTGTCTCTTTCTATTTCGAGCGGCTGGAAGGAGGCTCATGTGAGA 849  
 Db 727 GGCAAGCATGAGG-CTTGTCTCTTTCTATTTCGAGCGGCTGGAAGGAGGCTCATGTGAGA 785  
 QY 850 CTGTACCCATCCAGCGCTGCAATCAACATTTACTTAAACAGAGTCAACCGCAAC 909  
 Db 786 CTG-GTCCATCCAGCGCTGGGCTTCCCACTTTCTTTCAGGACAGCGCGCGCTGC 844  
 QY 910 ATGCTGTGTG 919  
 Db 845 TGGGGCTGGG 854

RESULT 13

LOCUS BG763582

DEFINITION 602735987F1 NIH\_MGC\_49 Homo sapiens cDNA clone IMAGE:4861268 5',  
 mRNA sequence.

ACCESSION BG763582

VERSION BG763582.1

KEYWORDS EST.

602735987F1 NIH\_MGC\_49 Homo sapiens cDNA clone IMAGE:4861268 5',  
 mRNA sequence.

602735987F1 NIH\_MGC\_49 Homo sapiens cDNA clone IMAGE:4861268 5',  
 mRNA sequence.

```

SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
REFERENCE    1. (bases 1 to 958)
AUTHORS     NIH-MGC http://mgc.nci.nih.gov/
TITLE       National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL     Unpublished
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgabbs-remail.nih.gov
            Tissue Procurement: ATCC/DCTD/DTG
            cDNA Library Preparation: Ling Hong/Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: L1CM1719 row: n column: 21
            High quality sequence stop: 810.

FEATURES    source
            Location/Qualifiers
            1..958
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="IMAGE:4861268"
            /tissue_type="melanotic melanoma, high MDR (cell line)"
            /lab_host="DH10B (phage-resistant)"
            /clone_lib="NIH MGC 49"
            /note="Organ: skin; Vector: pOTB7; Site 1: XhoI; Site 2:
            EcoRI; cDNA made by oligo-dT priming. Directionally cloned
            into EcoRI/XhoI sites using the following 5' adaptor:
            GGCACGAG(G). Size-selected >500bp for average insert size
            1.8kb. Library constructed by Ling Hong in the laboratory
            of Gerald M. Rubin (University of California, Berkeley)
            using ZAP-cDNA synthesis kit (Stratagene) and Superscript
            II RT (Life Technologies). Note: this is a NIH_MGC
            Library."

BASE COUNT  198 a 285 c 282 g 193 t
ORIGIN

Query Match      62.1%; Score 700.4; DB 12; Length 958;
Best Local Similarity 93.8%; Pred. No. 1.5e-171;
Matches 840; Conservative 0; Mismatches 41; Indels 15; Gaps 10;

QY 56 CCCAGCGCGCTCTGAGGAAACAGTACTGTCTACTTTGGGAATGGGTACGCTACCGTG 115
    |||
Db 61 CCCCTGCTCTCTGAGGAAACAGTACTGTCTACTTTGGGAATGGGTACGCTACCGTG 120
    |||

QY 116 GCACGCACAGCTCACCGAGTCGGTGCTCTCTGCTCCCGTGGAAATTCATGATCTCTGA 175
    |||
Db 121 GCACGCACAGCTCACCGAGTCGGTGCTCTCTGCTCCCGTGGAAATTCATGATCTCTGA 180
    |||

QY 176 TAGGCAAGTGTACACACACAGACACCCAGTCCCGAGCACTGGCGCTGGGCAACATA 235
    |||
Db 181 TAGGCAAGTGTACACACACAGACACCCAGTCCCGAGCACTGGCGCTGGGCAACATA 240
    |||

QY 236 ATTACTGCGGAATCCTGATGGGATGCCAAGCCCTGGTGGCAGCTGCTGAAGAACCGCA 295
    |||
Db 241 ATTACTGCGGATCTCTGATGGGATGCCAAGCCCTGGTGGCAGCTGCTGAAGAACCGCA 300
    |||

QY 296 GGCTGACGTGGAGTACTGTGATGTGCTCTCTGCTCCACTGGCGGCTGAGACAGTACA 355
    |||
Db 301 GGCTGACGTGGAGTACTGTGATGTGCTCTCTGCTCCACTGGCGGCTGAGACAGTACA 360
    |||

QY 356 GCCAGCCTCAGTTTCGATCAAGAGGGCTTTTCGCGGATGCGCTCCACCCCTGGC 415
    |||
Db 361 GCCAGCCTCAGTTTCGATCAAGAGGGCTTTTCGCGGATGCGCTCCACCCCTGGC 420
    |||

QY 416 AGGCTGCCATCTTTGCCAAGCACAGGAGGTCCCGCGGAGAGCGGTTCCTGTGCGGGGCA 475
    |||
Db 421 AGGCTGCCATCTTTGCCAAGCACAGGAGGTCCCGCGGAGAGCGGTTCCTGTGCGGGGCA 480
    |||

QY 476 TACTCATCAGTCTCTGCTGGATTTCTCTCTGCGGCCCACTGCTTCCAGGAGAGGTTTCCGC 535
    |||

```

```

Db 481 TACTCATCAGTCTCTGCTGGATTTCTCTCTCGGCCCACTGCTTCCAGGAGAGGTTTCTGT 540
    |||
QY 536 CCACACACCTGACGGTGCATCTTGGGGAGAAACATACCGGGTGGTCCCTTGGCGAGGAGGAGC 595
    |||
Db 541 CCACACACCTGACGGTGCATCTTGGGGAGAAACATACCGGGTGGTCCCTTGGCGAGGAGGAGC 600
    |||
QY 596 AGAAATTTGAAGTCGAAAAATACATTTCTCCATAAGGAATTCGATGATGACACTTACGACA 655
    |||
Db 601 AGAAATTTGAAGTCGAAAAATACATTTCTCCATAAGGAATTCGATGATGACACTTACGACA 660
    |||
QY 656 ATGACATTTGGCTGCTGACGCTGAAATCGAATTCGT--CCGCTGTGCCAGGAGAGCAGC 714
    |||
Db 661 ATGACATTTGGCTGCTGACGCTGAAATCGAATTCGTCCCGCTGTGCCAGGAGAGCAGC 720
    |||
QY 715 -GTGGTCCGACACTGTGTGCTT--CCCGCGCGGACCTGCAGCTGCCGAGCTGAC--GGA 770
    |||
Db 721 GTGGTCCGACACTGTGTGCTTCCCGCGCGGACCTGCAGCTGCCGAGCTGACCGGGA 780
    |||
QY 771 GTGTGAGCTCTCCG-GCTACGGCAAGCAT--GAGGCTTGTCTCTTTTATTCGAGAGCG 827
    |||
Db 781 GTGTGAGCTCTCCGAGCTACGGCAAGCATTTGAGGGCTGTGCTCTTCTTATTCGAGAGCG 840
    |||
QY 828 GCTGAAGGAGCTCATGTCTGACGTGTACCCATCCAGCGGCTGCACATCACAACTTTACT 887
    |||
Db 841 G-TGAAGGAGGCTCATG--CGACTGGACCCATTTCCACCGGTGGCCATTTCACAACTTTACT 897
    |||
QY 888 TAACAGAAACAGTACCACCAACATGCTGTGCTGGAGACATCTCGGAGCGCGCGGC 943
    |||
Db 898 T-ACAGAAACAGTACC---GAAATGCTGTGCGGACCCCGAGCGGGGGCCCGCAGC 949
    |||

```

```

RESULT 14
BO960216
LOCUS    BO960216
DEFINITION BO960216.1 GI:22375694
          5', mRNA sequence.
ACCESSION BO960216
VERSION    BO960216.1
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens

```

```

REFERENCE 1 (bases 1 to 974)
AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished
COMMENT   Contact: Robert Strausberg, Ph.D.
          Email: cgabbs-remail.nih.gov
          Tissue Procurement: ATCC
          cDNA Library Preparation: Rubin Laboratory
          cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
          DNA Sequencing by: Agencourt Bioscience Corporation
          Clone distribution: MGC clone distribution information can be
          found through the I.M.A.G.E. Consortium/LLNL at:
          http://image.llnl.gov
          Plate: L1CM2645 row: a column: 06
          High quality sequence stop: 603.

```

```

FEATURES    source
            Location/Qualifiers
            1..974
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="IMAGE:6459629"
            /tissue_type="epidermoid carcinoma, cell line"
            /lab_host="DH10B (phage-resistant)"
            /clone_lib="NIH MGC 101"
            /note="Organ: lung; Vector: pOTB7; Site 1: EcoRI; Site 2:
            XhoI; cDNA made by oligo-dT priming. Directionally cloned
            into EcoRI/XhoI sites using the following 5' adaptor:
            GGCACGAG(G). Library constructed by Ling Hong in the
            laboratory of Gerald M. Rubin (University of California,

```

Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH-MGC Library."

```
BASE COUNT      215 a   293 c   279 g   187 t
ORIGIN
Query Match      61.6%; Score 694.4; DB 13; Length 974;
Best Local Similarity 97.8%; Pred. No. 5.5e-170;
Matches 757; Conservative 0; Mismatches 11; Indels 6; Gaps 5;

QY 342 CCTGAGACAGTACAGCAGCCTCAGTTTCGATCAATCAAGAGGGCTCTTCGCCGACATCGC 401
Db 19  CCTGAGACAGTACAGCAGCCTCAGTTTCGATCAATCAAGAGGGCTCTTCGCCGACATCGC 78
QY 402 CTCACACCCCTGGCAGGCTGCATCTTTCGCAAGCAGAGGAGTCCGCCGAGAGCGGTT 461
Db 79  CTCACACCCCTGGCAGGCTGCATCTTTCGCAAGCAGAGGAGTCCGCCGAGAGCGGTT 138
QY 462 CTTGTGCGGGGGCATACTCATCAGCTCCTGTGGAATTTCTCTGCGGCCCACTGCTTCCA 521
Db 139 CTTGTGCGGGGGCATACTCATCAGCTCCTGTGGAATTTCTCTGCGGGCTC-TGCTTCCA 197
QY 522 GGAGAGGTTTCCGCCCAACCACTGACCGGTGATCTTGGCGAAGACATACCGGGTGGTCCC 581
Db 198 GGAGAGGTTTCCGCCCAACCACTGACCGGTGATCTTGGCGAAGACATACCGGGTGGTCCC 257
QY 582 TGGCGAGGAGGAGCAGAAATTTGAAGTCGAAATACATTTCCATAGGAATTCGATGA 641
Db 258 TGGCGAGGAGGAGCAGAAATTTGAAGTCGAAATACATTTCCATAGGAATTCGATGA 317
QY 642 TGACACTTACGACATGATTCGCTGCTGAGCTGAAATCGGATTCCTCCCGCTGTGC 701
Db 318 TGACACTTACGACATGATTCGCTGCTGAGCTGAAATCGGATTCCTCCCGCTGTGC 377
QY 702 CCAGAGAGCAGCTGTGTCGACATGTGTGCTTCCCGCGGAGCCTGCAGCTGCCGGA 761
Db 378 CCAGAGAGCAGCTGTGTCGACATGTGTGCTTCCCGCGGAGCCTGCAGCTGCCGGA 437
QY 762 CTGGAGCGAGTGTGAGCTCTCGGCTAGCGAAGCATCAGAGCCTGTCTCTTCTATTTC 821
Db 438 CTGGAGCGAGTGTGAGCTCTCGGCTAGCGAAGCATCAGAGCCTGTCTCTTCTATTTC 497
QY 822 GGAGCGGCTGAAGAGGCTCATGTGACAGCTGTACCCATCCAGCGGCTGCACATCACACA 881
Db 498 GGAGCGGCTGAAGAGGCTCATGTGACAGCTGTACCCATCCAGCGGCTGCACATCACACA 557
QY 882 TTTACTTAACAGAACAGTCACCGAACAATGCTGTGTGTGAGACATCTCGAGCGCGG 941
Db 558 TTTACTTAACAGAACAGTCACCGAACAATGCTGTGTGTGAGACATCTCGAGCGCGG 617
QY 942 GCCCCAGGCAACTTGCAGCGGCTGCCGAGGCGATTCGGGAGG-CGCCCTGTGTGTC 1000
Db 618 GCCCCAGGCAACTTGCAGCGGCTGCCGAGGCGATTCGGGAGGCCCCCTGTGTGTC 677
QY 1001 TGAACGATGGCGCAT-GACTTTGTTGGGCATCATCAGCT-GGGGCGCTGGGCTGTGGACA 1058
Db 678 TGAACGATGGCGCATGACCTTTGTTGGGCATCATCAGCTGGGCGCTGGGCTGTGGACA 737
QY 1059 GAAGATGTCCC--GGGTGTGTACAAAGGTTTACCACTACCTAGACTGGATT 1110
Db 738 GAAGATGTCCC--GGGTGTGTGTACCAAGGTTTACCACTACCTAAACTGGAAT 791
```

```
RESULT 15
BE732704
LOCUS      601571366F1 NIH_MGC_21 Homo sapiens cDNA clone IMAGE:3925600 5',
DEFINITION mRNA sequence.
ACCESSION BE732704
VERSION    BE732704.1 GI:10146696
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 709)  
NIH-MGC http://mgs.nci.nih.gov/.

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

Contact: Robert Strausberg, Ph.D.

Email: cga@nci.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: Ling Hong/Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov

Plate: L1CM752 row: d column: 17

High quality sequence start: 3

High quality sequence stop: 701.

FEATURES  
source

1. 709

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:3925600"

/tissue\_type="choriocarcinoma"

/lab\_host="DH10B (phage-resistant)"

/clone\_lib="NIH MGC 21"

/notes="Organ: placenta; Vector: pOTB7; Site: 1: XhoI;

Site: 2: EcoRI; cDNA made by oligo-dT priming.

Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by

Ling Hong in the laboratory of Gerald M. Rubin (University

of California, Berkeley) using ZAP-cDNA synthesis kit

(Stratagene) and Superscript II RT (Life Technologies)."

BASE COUNT 147 a 219 c 203 g 140 t

ORIGIN

Query Match 61.5%; Score 693.4; DB 10; Length 709;

Best Local Similarity 99.7%; Pred. No. 8.9e-170;

Matches 705; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 263 CCAAGCCCTGGTCCACGTCCTGAAGAACCGCAGGCTGACGTGGAGTACTGTGATGTC 322

Db 1 CCAAGCCCTGGTCCACGTCCTGAAGAACCGCAGGCTGACGTGGAGTACTGTGATGTC 60

QY 323 CTTCTGCTCCACTGGGCTGAGACAGTACAGCCAGCTCAGTTTCGATCAAGGAG 382

Db 61 CTTCTGCTCCACTGGGCTGAGACAGTACAGCCAGCTCAGTTTCGATCAAGGAG 120

QY 383 GGCTTTTCGCCGACATCGCTCCACCCCTGGCAGGCTGCCATCTTTGCCAAGCACAGGA 442

Db 121 GGCTTTTCGCCGACATCGCTCCACCCCTGGCAGGCTGCCATCTTTGCCAAGCACAGGA 180

QY 443 GGTGCCCCGAGAGCGGTTCTCTGTGGGGGGCATATCATCAGCTCTCTGTGGATTCTCT 502

Db 181 GGTGCCCCGAGAGCGGTTCTCTGTGGGGGGCATATCATCAGCTCTCTGTGGATTCTCT 240

QY 503 CTGCCCCCCTACTGTTCCAGAGAGGTTTCGGCCCCCACCACCTGACCGGTGATCTTGGGCA 562

Db 241 CTGCCCCCCTACTGTTCCAGAGAGGTTTCGGCCCCCACCACCTGACCGGTGATCTTGGGCA 300

QY 563 GAACATACCGGGTGGTCCCTGGCGAGGAGGAGCAGAAATTTGAAGTCAAAATATACATTG 622

Db 301 GAACATACCGGGTGGTCCCTGGCGAGGAGGAGCAGAAATTTGAAGTCAAAATATACATTG 360

QY 623 TCCATAGGAATTCGATGATGACATTTACGACAAATGACATTCGCTGCTCAGCTGAAAT 682

Db 361 TCCATAGGAATTCGATGATGACATTTACGACAAATGACATTCGCTGCTCAGCTGAAAT 420

QY 683 CGGATTCGTCCTCGCTGTGCCAGAGAGCAGCTGTGTCGCCGACTGTGTGCTTCCCCCGG 742

Db 421 CGGATTCGTCCTCGCTGTGCCAGAGAGCAGCTGTGTCGCCGACTGTGTGCTTCCCCCGG 480

QY 743 CGGACCTGCAGCTGCCGACTCGACGGAGTGTGAGCTCTCCGGCTACCGCAAGCATGAGG 802  
Db |||||  
QY 481 CGGACCTGCAGCTGCCGACTCGACGGAGTGTGAGCTCTCCGGCTACCGCAAGCATGAGG 540  
Db |||||  
QY 803 CCTTGTCTCCTTTCTATTTCGGAGCGGCTGAAGGAGGCTCATGTCAAGACTGTACCCATCCA 862  
Db |||||  
QY 541 CCTTGTCCCCCTTTCTATTTCGGAGCGGCTGAAGGAGGCTCATGTCAAGACTGTACCCATCCA 600  
Db |||||  
QY 863 GCCGCTGCACATCAACAATTACTTAAACAGAACAGTCAACGACAAACATGCTGTGTGCTG 922  
Db |||||  
QY 601 GCCGCTGCACATCAACAATTACTTAAACAGAACAGTCAACGACAAACATGCTGTGTGCTG 660  
Db |||||  
QY 923 GAGACACTCGGAGCGGCGGCCCGAGGCAAACTTGCACGACGCTGC 969  
Db |||||  
QY 661 GAGACACTCGGAGCGGCGGG-CCGAGGCAAACTTGCACGACGCTGC 706  
Db |||||

Search completed: August 19, 2003, 08:29:18  
Job time : 3300.95 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 18, 2003, 19:53:48 ; Search time 4736.31 Seconds  
(without alignments)  
9743.047 Million cell updates/sec

Title: US-09-987-455-2  
Perfect score: 1128  
Sequence: 1 atgaaagacagctatcgc.....ttcgtgacaacatgcgacccg 1128

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl.\*

1: gb.ba.\*

2: gb.htg.\*

3: gb.in.\*

4: gb.om.\*

5: gb.ov.\*

6: gb.pat.\*

7: gb.ph.\*

8: gb.pl.\*

9: gb.pr.\*

10: gb.ro.\*

11: gb.sts.\*

12: gb.by.\*

13: gb.un.\*

14: gb.vi.\*

15: em.ba.\*

16: em.fun.\*

17: em.hum.\*

18: em.in.\*

19: em.mu.\*

20: em.om.\*

21: em.or.\*

22: em.ov.\*

23: em.pat.\*

24: em.ph.\*

25: em.pl.\*

26: em.ro.\*

27: em.sts.\*

28: em.un.\*

29: em.vi.\*

30: em.htg.hum.\*

31: em.htg.inv.\*

32: em.htg.other.\*

33: em.htg.mus.\*

34: em.htg.pln.\*

35: em.htg.rtd.\*

36: em.htg.mam.\*

37: em.htg.vrt.\*

38: em.sv.\*

39: em.htgo.hum.\*

40: em.htgo.mus.\*

41: em.htgo.other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1128	100.0	1128	6	AX431268
2	1128	100.0	1128	6	AX431268 Sequence
3	1065	94.4	1314	6	AX431271
4	1065	94.4	1314	6	A27435 DNA sequenc
5	1065	94.4	1341	6	E01937
6	1065	94.4	1419	6	A30593 DNA for tis
7	1065	94.4	1419	6	A27725 DNA sequenc
8	1065	94.4	1419	6	AR059996
9	1065	94.4	1419	6	E01944
10	1065	94.4	1689	6	E06609
11	1065	94.4	1689	6	106614 Sequence 46
12	1065	94.4	1780	6	108789 Sequence 3
13	1065	94.4	1780	6	A06611
14	1065	94.4	1836	6	A10226
15	1065	94.4	1836	6	E01176
16	1065	94.4	1981	6	101581
17	1065	94.4	1989	6	101583
18	1065	94.4	2099	6	101256
19	1065	94.4	2100	6	A07207
20	1065	94.4	2100	6	A07197
21	1065	94.4	2100	6	A27429
22	1065	94.4	2101	6	E01934
23	1065	94.4	2151	6	AR059986
24	1065	94.4	2162	6	E00654
25	1065	94.4	2162	6	AR017909
26	1065	94.4	2162	6	AR017931
27	1065	94.4	2249	6	118796
28	1065	94.4	2544	6	A07209
29	1065	94.4	2544	6	108530
30	1065	94.4	2547	6	E01163
31	1065	94.4	2547	6	107841
32	1065	94.4	2547	6	107989
33	1065	94.4	2549	6	109622
34	1065	94.4	2560	6	E01055
35	1065	94.4	7533	6	A01465
36	1064	94.3	1689	6	108637
37	1063.4	94.3	1314	6	AR059989
38	1063.4	94.3	1506	6	E02247
39	1063.4	94.3	1689	6	A14927
40	1063.4	94.3	1689	6	A14930
41	1063.4	94.3	1689	6	A34204
42	1063.4	94.3	1689	6	A34205
43	1063.4	94.3	1689	6	AX427761
44	1063.4	94.3	1689	6	E02448
45	1063.4	94.3	1689	6	E02663

ALIGNMENTS

RESULT 1	AX431268	Sequence 2 from Patent WO0240650.	1128 bp	DNA	linear	PAT 28-JUN-2002
LOCUS	AX431268					
DEFINITION	AX431268					
ACCESSION	AX431268					
VERSION	AX431268.1	GI:21656150				
KEYWORDS						
SOURCE						
ORGANISM						
REFERENCE						
AUTHORS						
TITLE						
JOURNAL						



Db CCTCAGTTTCGATCAAGAGGGGCTCTTCCGCGACATCGCTCCACCCCTGGAGGCT 420  
Qy GCCATCTTTGGCAAGCAGGAGGTGCGCCGAGAGCGGTTCCTGTGGGGGCGATATCTC 480  
Db GCCATCTTTGGCAAGCAGGAGGTGCGCCGAGAGCGGTTCCTGTGGGGGCGATATCTC 480  
Qy ATCAGCTCTCTGTGATTTCTCTGTGCGCCCACTGTCTTCCAGAGAGAGTTTCCGCGCCAC 540  
Db ATCAGCTCTCTGTGATTTCTCTGTGCGCCCACTGTCTTCCAGAGAGAGTTTCCGCGCCAC 540  
Qy CACCTGACGGTGATCTTGGGAGAGACATACCGGGTGGTCCCTGGCGAGAGAGCAGAAA 600  
Db CACCTGACGGTGATCTTGGGAGAGACATACCGGGTGGTCCCTGGCGAGAGAGCAGAAA 600  
Qy TTTGAAGTCGAAAATACATTTGTCCATAAGGAATTCGATGATGACATTTACGACAAATGAC 660  
Db TTTGAAGTCGAAAATACATTTGTCCATAAGGAATTCGATGATGACATTTACGACAAATGAC 660  
Qy ATTGGCTGTGCGAGCTGAATTCGATTCGTTCCCGCTGTGCGCGAGAGCAGCGTGTGTC 720  
Db ATTGGCTGTGCGAGCTGAATTCGATTCGTTCCCGCTGTGCGCGAGAGCAGCGTGTGTC 720  
Qy CGCAGTGTGCTTCCCGCGGAGCCTGAGCTGCGGACTGCGGAGCTGAGAGTGTGAGTTC 780  
Db CGCAGTGTGCTTCCCGCGGAGCCTGAGCTGCGGACTGCGGAGCTGAGAGTGTGAGTTC 780  
Qy TCCGGCTACGGCAGCATGAGGCTTGTCTCTTCTATTCCGAGCGGCTGAAGAGGCT 840  
Db TCCGGCTACGGCAGCATGAGGCTTGTCTCTTCTATTCCGAGCGGCTGAAGAGGCT 840  
Qy CATGTGAGACTGTACCCATCCAGCGGTGACATCAACAATTTACTTTAAACAGAACATC 900  
Db CATGTGAGACTGTACCCATCCAGCGGTGACATCAACAATTTACTTTAAACAGAACATC 900  
Qy ACCGACACATGCTGTGTGCTGGAGACATCGAGCGCGGCGCCCGAGCAACTTGCAC 960  
Db ACCGACACATGCTGTGTGCTGGAGACATCGAGCGCGGCGCCCGAGCAACTTGCAC 960  
Qy GACGCTGCCAGGGCGATTTGGGAGGCGCCCTGGTGTGCTGAACGATGGCGCGATGACT 1020  
Db GACGCTGCCAGGGCGATTTGGGAGGCGCCCTGGTGTGCTGAACGATGGCGCGATGACT 1020  
Qy TTGTGGGCATCATCAGCTGGGGCTGGGTGTGGACAGAGAGATGTCCCGGTGTGPAC 1080  
Db TTGTGGGCATCATCAGCTGGGGCTGGGTGTGGACAGAGAGATGTCCCGGTGTGTAC 1080  
Qy ACAAGGTTACCACTACTAGCTGAGTTCGTGACACATGCGACCG 1128  
Db ACAAGGTTACCACTACTAGCTGAGTTCGTGACACATGCGACCG 1128

## RESULT 3

A27435 1314 bp DNA linear PAT 09-JUL-2002  
LOCUS A27435  
DEFINITION DNA sequence of coding region in pTOKPAdeltatp.  
ACCESSION A27435  
VERSION A27435.1 GI:21727230  
KEYWORDS  
SOURCE synthetic construct  
ORGANISM synthetic construct  
REFERENCE 1 (bases 1 to 1314)  
AUTHORS Niwa,M., Saito,Y., Sasaki,H., Hayashi,M., Notani,J. and Kobayashi,M.  
TITLE New tissue plasminogen activator  
JOURNAL Patent: EP 0302456-A 35 08-FEB-1989;  
FUJISAWA PHARMACEUTICAL CO., LTD  
FEATURES  
Location/Qualifiers  
1..1314  
/organism="synthetic construct"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:32630"  
BASE COUNT 286 a 386 c 393 g 249 t

## ORIGIN

Query Match 94.4%; Score 1065; DB 6; Length 1314;  
Best Local Similarity 99.5%; Pred. No. 6.5e-244;  
Matches 1068; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
Qy 56 CCCAGGGGGCTCTGAGGGAAAACAGTACTGTCTACTTTTGGGAATGGGTGACCCCTACCGTG 115  
Db 239 CCCCTGCTCTCTGAGGGAAAACAGTACTGTCTACTTTTGGGAATGGGTGACCCCTACCGTG 298  
Qy 116 GCAGGCACAGCCTCACCGAGTCCGGTGCCTCTCTGCTCCCGTGAATTCATGATCTCTGA 175  
Db 299 GCAGGCACAGCCTCACCGAGTCCGGTGCCTCTCTGCTCCCGTGAATTCATGATCTCTGA 358  
Qy 176 TAGCAAGAGTTTACACAGCACAGAACCCAGTGCCTCCAGGCACCTGGGCTGGGCAACATA 235  
Db 359 TAGCAAGAGTTTACACAGCACAGAACCCAGTGCCTCCAGGCACCTGGGCTGGGCAACATA 418  
Qy 236 ATTACTCCCGGAATCTCTGATGGGGATGCCAAGCCCTGGTGCACGCTGCTGAAGAACCGCA 295  
Db 419 ATTACTCCCGGAATCTCTGATGGGGATGCCAAGCCCTGGTGCACGCTGCTGAAGAACCGCA 478  
Qy 296 GGCTGAGTGGGAGTACTGTGATGTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 355  
Db 479 GGCTGAGTGGGAGTACTGTGATGTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 538  
Qy 356 GCCAGCCTCAGTTTCGCATCAAGAGGAGGCTCTTTCGCGGACATCGCTCCACCCCTGGC 415  
Db 539 GCCAGCCTCAGTTTCGCATCAAGAGGAGGCTCTTTCGCGGACATCGCTCCACCCCTGGC 598  
Qy 416 AGGTGTCATCTTTGCGCAAGCAGGAGTGCCTCCGAGAGCGGTTCCTGTGCGGGGGCA 475  
Db 599 AGGTGTCATCTTTGCGCAAGCAGGAGTGCCTCCGAGAGCGGTTCCTGTGCGGGGGCA 658  
Qy 476 TACTCATCAGTCTCTGTCTGATTTCTCTCTGCGGCCACTGCTTCCAGGAGAGTTTCCGC 535  
Db 659 TACTCATCAGTCTCTGTCTGATTTCTCTCTGCGGCCACTGCTTCCAGGAGAGTTTCCGC 718  
Qy 536 CCACACACCTGACGCTGATCTTGGGCGAACAATACCGGGTGGTCCCTGGCGAGAGGAGC 595  
Db 719 CCACACACCTGACGCTGATCTTGGGCGAACAATACCGGGTGGTCCCTGGCGAGGAGAGC 778  
Qy 596 AGAAATTTGAAGTCGAAAAATACATTTCTCCATAAGGAATTCGATGATGACATTTACGACA 655  
Db 779 AGAAATTTGAAGTCGAAAAATACATTTCTCCATAAGGAATTCGATGATGACATTTACGACA 838  
Qy 656 ATGACATTTGGCTCTGCTGACGTGAATTCGATTCGCTCCCGTGTGCCCGAGAGAGCAGCG 715  
Db 839 ATGACATTTGGCTCTGCTGACGTGAATTCGATTCGCTCCCGTGTGCCCGAGAGAGCAGCG 898  
Qy 716 TGGTCCGCACCTGTGTGCTTCCCGCGGACCTGCGAGCTGCCGGAATGGACGAGGTGTG 775  
Db 899 TGGTCCGCACCTGTGTGCTTCCCGCGGAGACCTGCGAGCTGCCGGAATGGACGAGGTGTG 958  
Qy 776 AGCTCTCCGGCTACGGCAAGCAGTGAAGCTTGTCTCTCTCTCTCTCTCTCTCTCTCTCT 835  
Db 959 AGCTCTCCGGCTACGGCAAGCAGTGAAGCTTGTCTCTCTCTCTCTCTCTCTCTCTCTCT 1018  
Qy 836 AGGCTCATGTGACACTGTACCCATCCAGCGCTGCACATCAACAATTTACTTTAACAGAA 895  
Db 1019 AGGCTCATGTGACACTGTACCCATCCAGCGCTGCACATCAACAATTTACTTTAACAGAA 1078  
Qy 896 CAGTCACCGCAACATGCTGTGCTGCGAGACACTCGGAGCGGGGCGCCCGAGCAAACT 955  
Db 1079 CAGTCACCGCAACATGCTGTGCTGCGAGACACTCGGAGCGGGGCGCCCGAGCAAACT 1138  
Qy 956 TGCAACGAGCGCTGCGAGGCGAATTCGGAGGCGCCCTCTGGTGTGTCTGAACGATGCGCGCA 1015  
Db 1139 TGCAACGAGCGCTGCGAGGCGAATTCGGAGGCGCCCTCTGGTGTGTCTGAACGATGCGCGCA 1198  
Qy 1016 TGACTTTGGTGGGCATCATCAGCTGGGGCTTGGGCTGTGGACAGAGATGTCGCGGGTG 1075  
Db 1199 TGACTTTGGTGGGCATCATCAGCTGGGGCTTGGGCTGTGGACAGAGATGTCGCGGGTG 1258



```
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
BASE COUNT      293 a      397 c      399 g      252 t
ORIGIN

Query Match      94.4%; Score 1065; DB 6; Length 1341;
Best Local Similarity 99.5%; Pred. No. 6.5e-244;
Matches 1068; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 56 CCCAGCGCGCTCTCAGGGAACAGTACTGCTACTTTGGGAATGGGTACGCTACCGTG 115
   |||
Db 269 CCCCTGCTCTCTCAGGGAACAGTACTGCTACTTTGGGAATGGGTACGCTACCGTG 328

QY 116 GCACGACAGCCTCACCGAGTCGGTGCTCTCTCCCTCCGCTGGAAATCCATGATCTGA 175
   |||
Db 329 GCACGACAGCCTCACCGAGTCGGTGCTCTCTCCCTCCGCTGGAAATCCATGATCTGA 388

QY 176 TAGGCAAGGTTTACACGACACGAAACCCAGTCCGAGGCACTGGGCTGGGCAAAACATA 235
   |||
Db 389 TAGGCAAGGTTTACACGACACGAAACCCAGTCCGAGGCACTGGGCTGGGCAAAACATA 448

QY 236 ATTACTGCGGGAATCCTGATGGGATGCCAAGCCCTGCTGTCACCTGCTGAAGAACCGCA 295
   |||
Db 449 ATTACTGCGGGAATCCTGATGGGATGCCAAGCCCTGCTGTCACCTGCTGAAGAACCGCA 508

QY 296 GCGTGACGTGGGAGTACTGTGATGTCCTCTCTCCACCTGCCGCTGAGACAGTACA 355
   |||
Db 509 GCGTGACGTGGGAGTACTGTGATGTCCTCTCTCCACCTGCCGCTGAGACAGTACA 568

QY 356 GCCAGCCTCAGTTTCGATCAAGAGAGGCTCTTCGCGCAATCGCTCCCAACCCCTGGC 415
   |||
Db 569 GCCAGCCTCAGTTTCGATCAAGAGAGGCTCTTCGCGCAATCGCTCCCAACCCCTGGC 628

QY 416 AGGTGCCATCTTTGCCAGCACAGAGGTGCCCGGAGAGCGTCTGTGCGGGGCA 475
   |||
Db 629 AGGTGCCATCTTTGCCAAGCACAGAGGTGCCCGGAGAGCGTCTGTGCGGGGCA 688

QY 476 TACTCATCAGTCTCTGCTGGATTTCTCTGCGGCCCACTGCTTCCAGGAGAGTTTCGCG 535
   |||
Db 689 TACTCATCAGTCTCTGCTGGATTTCTCTGCGGCCCACTGCTTCCAGGAGAGTTTCGCG 748

QY 536 CCCACCACTGACGCTGATCTTGGGCAAAACATACCCGGTGGTCCCTGGCGAGGAGGAGC 595
   |||
Db 749 CCCACCACTGACGCTGATCTTGGGCAAAACATACCCGGTGGTCCCTGGCGAGGAGGAGC 808

QY 596 AGAATTTGAAGTCGAAATACATTTGTCATAGGAATTCGATGATGACACTTACGACA 655
   |||
Db 809 AGAATTTGAAGTCGAAATACATTTGTCATAGGAATTCGATGATGACACTTACGACA 868

QY 656 ATGACATTCGCTGCTGAGCTGAATCGGATTCGTCCTGCTGTCGCCAGGAGAGCAGCG 715
   |||
Db 859 ATGACATTCGCTGCTGAGCTGAATCGGATTCGTCCTGCTGTCGCCAGGAGAGCAGCG 928

QY 716 TGGTCCGCACTGTGTGCTTCCCGCGGACCTGACGCTGCGGACTGGACGAGTGCTG 775
   |||
Db 929 TGGTCCGCACTGTGTGCTTCCCGCGGACCTGACGCTGCGGACTGGACGAGTGCTG 988

QY 776 AGCTCTCGGCTACGGCAAGCATGAGGCTTGTCTCTTTCTATTCGAGGCGGCTGAAGG 835
   |||
Db 989 AGCTCTCGGCTACGGCAAGCATGAGGCTTGTCTCTTTCTATTCGAGGCGGCTGAAGG 1048

QY 836 AGGCTCATGTGACAGTGTACCATCCAGCCGCTGCACATCAACATTTACTTACAGAA 895
   |||
Db 1049 AGGCTCATGTGACAGTGTACCATCCAGCCGCTGCACATCAACATTTACTTACAGAA 1108

QY 896 CAGTCAACGACAACTGTGTGTGTGCTGGAGACACTCGGAGGCGGGGCCCGAGGCAAACT 955
   |||
Db 1109 CAGTCAACGACAACTGTGTGTGTGCTGGAGACACTCGGAGGCGGGGCCCGAGGCAAACT 1168

QY 956 TGCACGACGCTGCCAGGGGAGTTTCGGAGGCGCCCTGCTGTGTGTAACGATGGCGGCA 1015
   |||
Db 1169 TGCACGACGCTGCCAGGGGAGTTTCGGAGGCGCCCTGCTGTGTGTAACGATGGCGGCA 1228
```



```

REFERENCE 1 (bases 1 to 1419)
AUTHORS Niwa,M., Saico,Y., Sasaki,H., Hayashi,M., Notani,J. and Kobayashi,M.
TITLE NOVEL TISSUE PLASMINOGEN ACTIVATION FACTOR
JOURNAL Patent: JP 1989104167-A 11 21-APR-1989; FUJISAWA PHARMACEUT CO LTD
COMMENT OS Artificial gene
OC Artificial sequence; Genes.
PN JP 1989104167-A/11
PD 21-APR-1989
PF 01-AUG-1988 JP 1988192320
PR 03-AUG-1987 GB 87 8718298, 26-OCT-1987 GB 87 8725052, PR 13-NOV-1987 GB 87 8726683
PI NIWA MINEO, SAITO YOSHIMASA, SASAKI HITOSHI, HAYASHI MASAKO, NOTANI JOJI,
PI KOBAYASHI MASAKAZU
PC C12N9/64, C12N1/20, C12N5/00, C12N15/00//A61K37/54, (C12N9/64, PC C12R1:19),
PC (C12N9/64, C12R1:91);
CC strandedness: Double;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
CC *source: clone=pmTQk12;
FH Key Location/Qualifiers
FT CDS 1..1419
FT /product='new t-PA'.
FEATURES
source Location/Qualifiers
1..1419
/mol_type='genomic DNA'
/db_xref='taxon:32630'
BASE COUNT 309 a 413 c 426 g 271 t
ORIGIN
Query Match 94.4%; Score 1065; DB 6; Length 1419;
Best Local Similarity 99.5%; Pred. No. 6.5e-244;
Matches 1068; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
Qy 56 CCCAGGCGGCTCTGAGGAAACAGTACTCTCTTTGGGAATGGGTCAGCCTACCGTG 115
Db |||
Qy 344 CCCCTGCTCTGTAGGAAACAGTACTCTCTTTGGGAATGGGTCAGCCTACCGTG 403
Db |||
Qy 116 GCACGCACAGCCTCACCGAGTCGGTGCTCTCTCTCCCGTGGAAATTCATGATCTGA 175
Db |||
Qy 404 GCACGCACAGCCTCACCGAGTCGGTGCTCTCTCTCCCGTGGAAATTCATGATCTGA 463
Db |||
Qy 176 TAGGCAAGGTTTACACAGCAGACAGACCCAGTGGCCAGGCACTGGGCTGGGCAACATA 235
Db |||
Qy 464 TAGGCAAGGTTTACACAGCAGACAGACCCAGTGGCCAGGCACTGGGCTGGGCAACATA 523
Db |||
Qy 236 ATTACTGCGGGAATCTCTATGGGATGCCAAGCCTGTGTGTCACGTGCTGAAGAACGCA 295
Db |||
Qy 524 ATTACTGCGGGAATCTCTATGGGATGCCAAGCCTGTGTGTCACGTGCTGAAGAACGCA 583
Db |||
Qy 296 GGCTGACGTGGGAGTACTGTGATGTGCTCTCTCTCTCCACCTGGGCTGGAGACAGTACA 355
Db |||
Qy 584 GGCTGACGTGGGAGTACTGTGATGTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 643
Db |||
Qy 356 GCCAGCCTCAGTTTGCATCAAGAGAGGCTCTCTGCGCGCATCGCTCCACCCCTGGC 415
Db |||
Qy 644 GCCAGCCTCAGTTTGCATCAAGAGAGGCTCTCTGCGCGCATCGCTCCACCCCTGGC 703
Db |||
Qy 416 AGGCTGCCATCTTTGCCAAGCAGAGAGGTCGCGGAGAGGTTCTCTGTGCGGGGCA 475
Db |||
Qy 704 AGGCTGCCATCTTTGCCAAGCAGAGAGGTCGCGGAGAGGTTCTCTGTGCGGGGCA 763
Db |||
Qy 476 TACTCATCAGCTCTCTGTGGATTTCTCTCTGCGGCCACTGTCTTCAGGAGAGGTTTCGCG 535
Db |||
Qy 764 TACTCATCAGCTCTCTGTGGATTTCTCTCTGCGGCCACTGTCTTCAGGAGAGGTTTCGCG 823
Db |||
Qy 536 CCCACCCTGACGGTGTATCTTTGGGCAACAATACCGGGTGGTCTCTCTGGGAGGAGC 595
Db |||

Db 824 CCCACCCTGACGGTGTATCTTTGGGCAACAATACCGGGTGGTCTCTCTGGGAGGAGC 883
Qy |||
Qy 596 AGAAATTTGAAGTCGAAATATACATTTGTCCATAAGGAATTCGATGATGACACTTACGACA 655
Db |||
Qy 884 AGAAATTTGAAGTCGAAATATACATTTGTCCATAAGGAATTCGATGATGACACTTACGACA 943
Qy |||
Qy 656 ATGACATTTGGCTGTCTGAGCTGAAATCGAATTCGTCCTCCGCTGTGCCAGAGAGCAGCG 715
Db |||
Qy 944 ATGACATTTGGCTGTCTGAGCTGAAATCGAATTCGTCCTCCGCTGTGCCAGAGAGCAGCG 1003
Qy |||
Qy 716 TGGTCCGACCTGTGCTCTTCCCGCGGAGCCTGCGAGCTGCGGAGTGGACGAGTGTG 775
Db |||
Qy 1004 TGGTCCGACCTGTGCTCTTCCCGCGGAGCCTGCGAGCTGCGGAGTGGACGAGTGTG 1063
Qy |||
Qy 776 AGCTCTCTGGCTACGGCAAGCATGAGGCTTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 835
Db |||
Qy 1064 AGCTCTCTGGCTACGGCAAGCATGAGGCTTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1123
Qy |||
Qy 836 AGGCTCATGTGACACTGTATCCCATCCAGCCGCTGCATCATCAACAATTTACTTAAAGAA 895
Db |||
Qy 1124 AGGCTCATGTGACACTGTATCCCATCCAGCCGCTGCATCATCAACAATTTACTTAAAGAA 1183
Qy |||
Qy 896 CAGTCACGCAACATGCTGTGCTGAGACACATCTCGAGCGGCGGCGCCAGGCAACT 955
Db |||
Qy 1184 CAGTCACGCAACATGCTGTGCTGAGACACATCTCGAGCGGCGGCGCCAGGCAACT 1243
Qy |||
Qy 956 TGCAACAGCGCTGCCAGGCGATTTCCGAGAGCCCTCTGCTGTCTGTAACATGCGCGCA 1015
Db |||
Qy 1244 TGCAACAGCGCTGCCAGGCGATTTCCGAGAGCCCTCTGCTGTCTGTAACATGCGCGCA 1303
Qy |||
Qy 1016 TGACTTTGGTGGGCTCATCAGCTGGGCTGTGGAGAGAGATGTCCCGGTG 1075
Db |||
Qy 1304 TGACTTTGGTGGGCTCATCAGCTGGGCTGTGGAGAGAGATGTCCCGGTG 1363
Qy |||
Qy 1076 TGTACAAAGGTTTACCAACTCCTAGACTGGAATTCGTGACAACTGCGACG 1128
Db |||
Qy 1364 TGTACAAAGGTTTACCAACTCCTAGACTGGAATTCGTGACAACTGCGACG 1416
Db |||

RESULT 9
LOCUS I06609 1689 bp DNA linear PAT 02-DEC-1994
DEFINITION Sequence 44 from Patent EP 0293934.
ACCESSION I06609
VERSION I06609.1 GI:590639
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1689)
AUTHORS Mulvihill,E.R., Nexo,B.A., Yoshitake,S., Ikeda,Y., Suzuki,S., Hashimoto,A. and Yuzuriha,T.
TITLE Mutant t-PA with kringle replacement
JOURNAL Patent: EP 0293934-A1 44 07-DEC-1988;
FEATURES Location/Qualifiers
source 1..1689
/organism='unknown'
BASE COUNT 377 a 483 c 505 g 324 t
ORIGIN
Query Match 94.4%; Score 1065; DB 6; Length 1689;
Best Local Similarity 99.5%; Pred. No. 6.5e-244;
Matches 1068; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
Qy 56 CCCAGGCGGCTCTGAGGAAACAGTACTCTCTTTGGGAATGGGTCAGCCTACCGTG 115
Db |||
Qy 614 CCCCTGCTCTGTAGGAAACAGTACTCTCTTTGGGAATGGGTCAGCCTACCGTG 673
Db |||
Qy 116 GCACGCACAGCCTCACCGAGTCGGTGCTCTCTCTCCCGTGGAAATTCATGATCTGA 175
Db |||
Qy 674 GCACGCACAGCCTCACCGAGTCGGTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 733
Db |||

```







QY 896 CAGTCACCGACAACTGCTGTGTGCTGGAGACACTCGGAGCGCGCGGCCCAAGCAAACT 955  
 Db 1454 CAGTCACCGACAACTGCTGTGTGCTGGAGACACTCGGAGCGCGCGGCCCAAGCAAACT 1513  
 QY 956 TGCACGAGCGCTGCCAGGCGGATTCGGAGGCGCCCTGTGTGTCTGAACGATGCGCGCA 1015  
 Db 1514 TGCACGAGCGCTGCCAGGCGGATTCGGAGGCGCCCTGTGTGTCTGAACGATGCGCGCA 1573  
 QY 1016 TGACTTTGGTGGGACATCATCACTAGCTGGGCGCTGGGCTGTGGACAGAGATGTCCCGGGTG 1075  
 Db 1574 TGACTTTGGTGGGACATCATCACTAGCTGGGCGCTGGGCTGTGGACAGAGATGTCCCGGGTG 1633  
 QY 1076 TGTACAAAAGTTACCAACTACCTAGACTGGATTCTGTGACAAATGCGACCG 1128  
 Db 1634 TGTACAAAAGTTACCAACTACCTAGACTGGATTCTGTGACAAATGCGACCG 1686

## RESULT 11

108789 I08789 1689 bp DNA linear PAT 02-DEC-1994  
 LOCUS Sequence 3 from Patent WO 8804690.

ACCESSION I08789

VERSION I08789.1 GI:588494

KEYWORDS Unknown.

SOURCE Unknown.

ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 1689)

AUTHORS Bollen,A.J., Gheysen,D., Jacobs,P., Pierard,L. and Collen,D.J.

JOURNAL Patent: WO 8804690-A 3 30-JUN-1988;

FEATURES Location/Qualifiers

1..1689

/organism="unknown"

BASE COUNT 379 a 480 c 505 g 325 t

ORIGIN

Query Match 94.4%; Score 1065; DB 6; Length 1689;

Best Local Similarity 99.5%; Pred. No. 6.5e-244;

Matches 1068; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 56 CCCAGCGCGCTCTGAGGAAACAGTCTACTTTGGGAATGGGTACGCTACCGTG 115  
 Db 614 CCCCTGCTCTGAGGAAACAGTCTACTTTGGGAATGGGTACGCTACCGTG 673  
 QY 116 GCACGACAGCTCAACGAGTCGGTGCTCTGCTCCCGTGGAAATCCATGATCTGA 175  
 Db 674 GCACGACAGCTCAACGAGTCGGTGCTCTGCTCCCGTGGAAATCCATGATCTGA 733  
 QY 176 TAGGCAAGTTTACACAGACAGAACCCAGTGCAGGCACTGGGCTGGGCAAAACATA 235  
 Db 734 TAGGCAAGTTTACACAGACAGAACCCAGTGCAGGCACTGGGCTGGGCAAAACATA 793  
 QY 236 ATTACTGCGGAACTCTGATGGGATGCCAGCCCTGGTGCACGTGTGAAGACCGCA 295  
 Db 794 ATTACTGCGGAACTCTGATGGGATGCCAGCCCTGGTGCACGTGTGAAGACCGCA 853  
 QY 296 GGCTGACGTGGGAGTACTGTGATGTCCTCTGCTCCACCTGCGGCTGGAGACAGTACA 355  
 Db 854 GGCTGACGTGGGAGTACTGTGATGTCCTCTGCTCCACCTGCGGCTGGAGACAGTACA 913  
 QY 356 GCCAGCCTCAGTTTCGATCAAAAGAGGGCTCTTCGCGGACATCGCCTCCACCCCTGGC 415  
 Db 914 GCCAGCCTCAGTTTCGATCAAAAGAGGGCTCTTCGCGGACATCGCCTCCACCCCTGGC 973  
 QY 416 AGGCTGCCATCTTTGCCAGACAGGAGTGCAGGAGGGTTCCTGTGCGGGGCA 475  
 Db 974 AGGCTGCCATCTTTGCCAGACAGGAGTGCAGGAGGGTTCCTGTGCGGGGCA 1033  
 QY 476 TACTCATCAGCTCTGCTGGATTCTCTGTCGCGCCCACTGCTTCAGGAGAGTTTCGCG 535  
 Db 1034 TACTCATCAGCTCTGCTGGATTCTCTGTCGCGCCCACTGCTTCAGGAGAGTTTCGCG 1093  
 QY 536 CCCACCCTGACGGTGATCTTTGGGACAGAACATACCGGGTGTCTCCCTGGCGAGGAGGAGC 595

Db 1094 CCCACCCTGACGGTGATCTTTGGGACAGAAATACACCGGGTGTCTCCCTGGCGAGGAGGAGC 1153  
 QY 596 AGAAATTTGAAGTCGAAAAATACATTTGTCATAGGAATTCGATGATGACACTTACGACA 655  
 Db 1154 AGAAATTTGAAGTCGAAAAATACATTTGTCATAGGAATTCGATGATGACACTTACGACA 1213  
 QY 656 ATGACATTTGCGCTCTGCGAGCTGAAATCGGATTCGCTCCCGCTGTGCCAGAGAGACGCG 715  
 Db 1214 ATGACATTTGCGCTCTGCGAGCTGAAATCGGATTCGCTCCCGCTGTGCCAGAGAGACGCG 1273  
 QY 716 TGGTCCGACCTGTGTGCTTCCCGCGGAGACCTGCGAGCTGCGGACCTGCGAGAGAGTGTG 775  
 Db 1274 TGGTCCGACCTGTGTGCTTCCCGCGGAGACCTGCGAGCTGCGGACCTGCGAGAGTGTG 1333  
 QY 776 AGCTCTCCGGCTACGGCAAGCATGAGCCTTGTCTCTCTTTCTATTTCGAGCGGCTGAAGG 835  
 Db 1334 AGCTCTCCGGCTACGGCAAGCATGAGCCTTGTCTCTCTTTCTATTTCGAGCGGCTGAAGG 1393  
 QY 836 AGGCTCATGTGACACTGTACCCATCCAGCGCTGCACATCAACAATTTACTTAACAGAA 895  
 Db 1394 AGGCTCATGTGACACTGTACCCATCCAGCGCTGCACATCAACAATTTACTTAACAGAA 1453  
 QY 896 CAGTCACCGACAACTGCTGTGCTGCGAGACACTCGAGCGCGCGGCCCCAGGCAAACT 955  
 Db 1454 CAGTCACCGACAACTGCTGTGCTGCGAGACACTCGAGCGCGCGGCCCCAGGCAAACT 1513  
 QY 956 TGACAGAGCGCTGCCAGGCGATTCGGAGGCGCCCTGTGTGCTGAACGATGCGCGCA 1015  
 Db 1514 TGACAGAGCGCTGCCAGGCGATTCGGAGGCGCCCTGTGTGCTGAACGATGCGCGCA 1573  
 QY 1016 TGACTTTGGTGGGACATCATCAGCTGGGCGCTGGGCTGTGGACAGAGATGTCCCGGGTG 1075  
 Db 1574 TGACTTTGGTGGGACATCATCAGCTGGGCGCTGGGCTGTGGACAGAGATGTCCCGGGTG 1633  
 QY 1076 TGACACAAAGTTTACCAACTACCTAGACTGGATTCTGTGACAAATGCGACCG 1128  
 Db 1634 TGACACAAAGTTTACCAACTACCTAGACTGGATTCTGTGACAAATGCGACCG 1686

## RESULT 12

A06611

LOCUS

DEFINITION

pt-PABam.

ACCESSION

A06611

VERSION

A06611.1

KEYWORDS

synthetic construct

SOURCE

synthetic construct

ORGANISM

artificial sequences.

REFERENCE

1 (bases 1 to 1780)

AUTHORS

JOURNAL

FEATURES

source

Patent: WO 900600-A 11 25-JAN-1990;

Location/Qualifiers

1..1780

/organism="synthetic construct"

/mol\_type="genomic DNA"

/db\_xref="taxon:32630"

17..1705

/codon\_start=1

/product="t-PA analogue"

/protein\_id="CAA00577.1"

/db\_xref="GI:412907"

/translation="MDAMKRLGCLVLLCGAVFVPSQEIHARFRGARSQVTCRDE

KTOMIQHQSGLRPLNRVYWCNSGRACQCHSVFKSCSPFCNFGTCCQALY

FSDFVCCPSPGFAKCEIDTRATCYEDOGISYRGWTSTAESGAECTNMSSALAKP

NGSRPDAIRLGLGHNHCENPDKPCWCVFKAGKYSEFSTPACSGNSDCYPG

WCHVLRNRLTWBYCDVPSCTGLROYSQFRIGLGLFADTASHPWQAIPAKHRR

SPERFLCCGILLISSCWILSAHCFQRPFPPLHTVLTRTYRVVPEEESQKPEVKY

IVHKEFDDDDTDNDIALQLKSDSSRCQAQESSVVVRVTVCLPPADLQLPDWTECELSGVG



Db 630 GCACGACAGCCTCACCGAGTCGGGTGCTCTCTGCTCCCGTGGAAATTCATGATCCTGA 749  
QY 176 TAGGCAAGGTTTACACAGACAGAACCCAGTCCAGGCACTGGGCTGGGCAAAACATA 235  
Db 750 TAGGCAAGGTTTACACAGACAGAACCCAGTCCAGGCACTGGGCTGGGCAAAACATA 809  
QY 236 ATTACTGCGGATCTGATGGGATGCAAGCCCTGTGTGCACTGTGTAAGAACCGCA 295  
Db 810 ATTACTGCGGATCTGATGGGATGCAAGCCCTGTGTGCACTGTGTAAGAACCGCA 869  
QY 296 GGCCTGACGTGGGAGTACTGATGTGCTCTCTGCTCCACCTGCGGCTGGAGACAGTACA 355  
Db 870 GGCCTGACGTGGGAGTACTGATGTGCTCTCTGCTCCACCTGCGGCTGGAGACAGTACA 929  
QY 356 GCCAGCCTCAGTTTCGATCAAGAGAGGCTCTTCGCGACATCGCTCCACCCCTGGC 415  
Db 930 GCCAGCCTCAGTTTCGATCAAGAGAGGCTCTTCGCGACATCGCTCCACCCCTGGC 989  
QY 416 AGGCTGCCATCTTTGCCAAGCAGGAGGTGCGCGGAGAGGCTTCTGTGCGGGGCA 475  
Db 990 AGGCTGCCATCTTTGCCAAGCAGGAGGTGCGCGGAGAGGCTTCTGTGCGGGGCA 1049  
QY 476 TACTCATCAGCTCCGTGCTGATCTCTGCGGCCACTGCTCCAGGAGGTTTCCGC 535  
Db 1050 TACTCATCAGCTCCGTGCTGATCTCTGCGGCCACTGCTCCAGGAGGTTTCCGC 1109  
QY 536 CCCACCACCTGACGCTGATCTTTGGGAGAACATACCGGCTGCTCCGCGAGGAGGAGC 595  
Db 1110 CCCACCACCTGACGCTGATCTTTGGGAGAACATACCGGCTGCTCCGCGAGGAGGAGC 1169  
QY 596 AGAAATTTGAAGTCGAAATATCATTTGCTCATAGGAATTCATGATGACACTTACGACA 655  
Db 1170 AGAAATTTGAAGTCGAAATATCATTTGCTCATAGGAATTCATGATGACACTTACGACA 1229  
QY 656 ATGACATTCGCTGCTGACGCTGAATCGATTCGCTCCGCTGTGCCAGGAGGACGCG 715  
Db 1230 ATGACATTCGCTGCTGACGCTGAATCGATTCGCTCCGCTGTGCCAGGAGGACGCG 1289  
QY 716 TGGTCCGACATGTGTGCTTCCCGCGGACCTGACGCTGCGGACTGGAAGGAGTGTG 775  
Db 1290 TGGTCCGACATGTGTGCTTCCCGCGGACCTGACGCTGCGGACTGGAAGGAGTGTG 1349  
QY 776 AGCTCTCGGCTACCGCAAGCATGAGCTGTGTCTCTTTCTATTTCGAGCGGCTGAAGG 835  
Db 1350 AGCTCTCGGCTACCGCAAGCATGAGCTGTGTCTCTTTCTATTTCGAGCGGCTGAAGG 1409  
QY 836 AGGCTCATGTGACACTGTACCCATCCAGCGCTGCACATCAACATTTACTTACAGAA 895  
Db 1410 AGGCTCATGTGACACTGTACCCATCCAGCGCTGCACATCAACATTTACTTACAGAA 1469  
QY 896 CAGTCACCGACAAATGTGTGTGTGAGACACTCGGAGCGGGGCCCCAGGCAAACT 955  
Db 1470 CAGTCACCGACAAATGTGTGTGTGAGACACTCGGAGCGGGGCCCCAGGCAAACT 1529  
QY 956 TGCAGGACGCTGCCAGGCGGATTCGGAGGCGCCCTGGTGTGTGAACGATGCGCGCA 1015  
Db 1530 TGCAGGACGCTGCCAGGCGGATTCGGAGGCGCCCTGGTGTGTGAACGATGCGCGCA 1589  
QY 1016 TGACTTTGGTGGGATCATCAGCTGGGCGCTGGGCTGTGGACAGGAGTGTCCCGGTG 1075  
Db 1590 TGACTTTGGTGGGATCATCAGCTGGGCGCTGGGCTGTGGACAGGAGTGTCCCGGTG 1649  
QY 1076 TGTACAAAAGGTTACCAACTACCTAGACTGATTCGTCGACAAATGCGACCG 1128  
Db 1650 TGTACAAAAGGTTACCAACTACCTAGACTGATTCGTCGACAAATGCGACCG 1702

## RESULT 14

LOCUS E01176 1836 bp DNA linear PAT 29-SEP-1997  
DEFINITION DNA encoding tris-kringle plasminogen activator.  
ACCESSION E01176  
VERSION E01176.1 GI:2169435

## KEYWORDS

SOURCE synthetic construct

## ORGANISM

artificial construct

artificial sequences.

## REFERENCE

1 (bases 1 to 1836)

## AUTHORS

Pooru, P.H., Narendaa, K.K. and Shiyouguangu, R.R.

## TITLE

POLYKINGLE PLASMINOGEN ACTIVATING FACTOR

## JOURNAL

Patent: JP 1987104577-A 1 15-MAY-1987;

## COMMENT

AMERICAN HOME PROD CORP

OS Artificial gene

OC Artificial sequence; Genes.

PN JP 1987104577-A/1

PD 15-MAY-1987

PF 13-AUG-1986 JP 1986191173

PI 14-AUG-1985 US 85 766163, 11-JUL-1986 US 86 884835

RIN POORU POOUEN HANGU, NARENDAA KUMARU KARUYAN, PI SHIYOUGUANGU

R11 PC C12N9/52, A61K35/12, A61K35/74, A61K37/02, A61K37/54, C07H21/04, PC

C12N15/00,

CC (C12N9/52, C12R1:19), (C12N15/00, C12R1:19);

CC strandedness: Double;

CC topology: Linear;

CC hypothetical: No;

CC anti-sense: No;

CC \*source: cell type=fibroblast;

CC \*source: cell\_line=WI-38cell;

FH Key Location/Qualifiers

FT CDS 12..1793

FT /product='this-kringle plasminogen activator'

FT sig\_peptide 12..71

FT /product='signal peptide of urokinase' FT

mat\_peptide 72..1790

FT /product='this-kringle plasminogen activator'

FEATURES

source

1..1836

/organism='synthetic construct'

/mol\_type='genomic DNA'

/db\_xref='taxon:32630'

BASE COUNT 422 a 537 c 528 g 349 t

ORIGIN

Query Match 94.4%; Score 1065; DB 6; Length 1836;

Best Local Similarity 99.5%; Pred. No. 6.5e-244;

Matches 1068; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 56 CCCAGGCGGCTCTGAGGGAACAGTGTCTACTTTGGGAATGGGTACGCTTACCGTG 115

Db 718 CCCCTGCTGCTCTGAGGGAACAGTGTCTACTTTGGGAATGGGTACGCTTACCGTG 777

QY 116 GCAGCAGAGCTCACCGAGTCCGCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCT 175

Db 778 GCAGCAGAGCTCACCGAGTCCGCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCT 837

QY 176 TAGCAAGGTTTACACAGCAGAACCCAGTCCAGGCACTGGGCTGGGCAAAACATA 235

Db 838 TAGCAAGGTTTACACAGCAGAACCCAGTCCAGGCACTGGGCTGGGCAAAACATA 897

QY 236 ATTACTCCGGAATCTTGATGGGATGCCAAGCCCTGGTGCACGTGCTGAAGAACCGCA 295

Db 898 ATTACTCCGGAATCTTGATGGGATGCCAAGCCCTGGTGCACGTGCTGAAGAACCGCA 957

QY 296 GGCTGAGTGGGAGTACTGTGATGTGCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCT 355

Db 958 GGCTGAGTGGGAGTACTGTGATGTGCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCT 1017

QY 356 GCCAGCCTCAGTTTCGATCAAGAGGAGGCTCTTCGCGGACATCGCTCCACCCCTGGC 415

Db 1018 GCCAGCCTCAGTTTCGATCAAGAGGAGGCTCTTCGCGGACATCGCTCCACCCCTGGC 1077

QY 416 AGGCTGCCATCTTTGCAAGCAGAGGAGTGCAGGAGGAGGCTTCTGTGCGGGGCA 475

[illegible]

Search completed: August 19, 2003, 04:46:27

Job time : 4739.31 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 18, 2003, 19:51:53 ; Search time 397.386 Seconds  
(without alignments)  
7662.485 Million cell updates/sec

Title: US-09-987-455-2

Perfect score: 1128

Sequence: 1 atgaaaagacagctatgc.....ttctgacacatgcgacgc 1128

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N\_Geneseq 19Jun03.\*

- 1: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.\*
- 2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.\*
- 3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.\*
- 4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.\*
- 5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.\*
- 6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.\*
- 7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.\*
- 8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.\*
- 9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.\*
- 10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.\*
- 11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.\*
- 12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.\*
- 13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.\*
- 14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.\*
- 15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.\*
- 16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.\*
- 17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.\*
- 18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.\*
- 19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.\*
- 20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.\*
- 21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.\*
- 22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.\*
- 23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.\*
- 24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.\*
- 25: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1128	100.0	1128	24	AAD40614
2	1065	94.4	1314	10	AA091123
3	1065	94.4	1341	11	AA06762
4	1065	94.4	1419	10	AA091133
5	1065	94.4	1689	9	AA082178
6	1065	94.4	1780	11	AA001358
7	1065	94.4	2100	10	AA091119
8	1065	94.4	2100	11	AA05532

9	1065	94.4	2162	16	AA087370	Human tPA clone.
10	1065	94.4	2162	19	AA037294	Human tissue plasminogen
11	1065	94.4	2547	7	AA060659	Sequence encoding
12	1065	94.4	2560	10	AA090542	Plasmin pK12 cont
13	1065	94.4	7533	8	AA070491	Entire sequence of
14	1065	94.4	7533	9	AA081970	Plasmin pK12-tPA.
15	1064	94.3	1680	17	AA027588	Novel plasminogen
16	1064	94.3	1680	17	AA027587	Novel plasminogen
17	1063.4	94.3	1614	11	AA030581	Glycosylated plasmin
18	1063.4	94.3	1614	12	AA011550	Sequence encoding
19	1063.4	94.3	1659	12	AA011551	Sequence encoding
20	1063.4	94.3	1686	21	AA052893	Tissue type plasmin
21	1063.4	94.3	1689	9	AA082179	Tissue plasminogen
22	1063.4	94.3	1689	11	AA081090	Mutated t-PA analo
23	1063.4	94.3	1689	11	AA040903	Part of tPA024 gen
24	1063.4	94.3	1689	11	AA040904	Part of tPA023 gen
25	1063.4	94.3	1689	11	AA05177	Sequence encoding
26	1063.4	94.3	1689	12	AA013379	T-PA67+ mutant wit
27	1063.4	94.3	1689	24	AA052546	Human tissue plasmin
28	1063.4	94.3	1740	17	AA027585	Full-length tissue
29	1063.4	94.3	1955	21	AA021998	Human tissue plasmin
30	1063.4	94.3	1989	7	AA060700	Sequence encoding
31	1063.4	94.3	2091	12	AA012073	T-PA variant contg
32	1063.4	94.3	2099	11	AA05534	Plasmin pLNTQTPA
33	1063.4	94.3	2100	12	AA012072	T-PA with -ve char
34	1063.4	94.3	2170	8	AA070248	Sequence encoding
35	1063.4	94.3	2170	16	AA086576	Human tissue plasmin
36	1063.4	94.3	2249	11	AA05535	Plasmin pLGN delta
37	1063.4	94.3	2313	11	AA005944	Sequence encoding
38	1063.4	94.3	2457	7	AA060146	cDNA sequence enco
39	1063.4	94.3	2509	10	AA090508	cDNA of human tiss
40	1063.4	94.3	2519	24	AA052545	Human tissue plasmin
41	1063.4	94.3	2519	24	AA059544	Human tissue type
42	1063.4	94.3	2547	4	AA030001	Sequence encoding
43	1063.4	94.3	7360	17	AA015930	DHFR/intron (WT)ras
44	1062	94.1	1065	24	AA040616	Human kringie 2 se
45	1062	94.1	1068	10	AA091120	Sequence of coding

ALIGNMENTS

RESULT 1

AAD40614

ID AAD40614 standard; DNA; 1128 BP.

XX AAD40614;

AC AAD40614;

XX 30-OCT-2002 (first entry)

XX OmpA-K2S fusion protein encoding DNA.

DB Tissue plasminogen activator; tPA; kringie 2 serine protease; stroke;

KW K2S; cardiac infarction; acute myocardial infarction; pulmonary embolism;

KW artery occlusion; deep vein thrombosis; blood clotting; thrombolytic;

KW cerebroprotective; cardiant; ompA; fusion protein; gene; ds.

XX Unidentified.

OS Unidentified.

XX Key Location/Qualifiers

FT CDS 1..1128

FT /\*tag= a

FT /product= "OmpA-K2S fusion protein"

FT /transl\_except= (pos:1126..1128, aa:Pro-Gly)

FT /notes= "No stop codon"

FT /partial

FT WO200240650-A2.

XX 23-MAY-2002.

XX 07-NOV-2001; 2001WO-EPI2857.

XX





Db	1019	AGGTCATGTCAGACTGATCCCATCCAGCCGCTGCATCATCAAACTTTACTTAACGAA	1078
Qy	896	CAGTCACCGCAACAATGCTGTGCTCGAGACACTCGGAGCGCGGGCCCCAGGCAAACT	955
Db	1079	CAGTCACCGCAACAATGCTGTGCTCGAGACACTCGGAGCGCGGGCCCCAGGCAAACT	1138
Qy	956	TGCACGACGCTGCGAGGGCGATTCGGAGGCGCCCTGGTGTCTGAACGATGCGCCGCA	1015
Db	1139	TGCACGACGCTGCGAGGGCGATTCGGAGGCGCCCTGGTGTCTGAACGATGCGCCGCA	1199
Qy	1016	TGACTTTGGTGGGCATCATCAGCTGGGGCCCTGGGCTGTGACAGAGGATGTCCCGGGTG	1075
Db	1199	TGACTTTGGTGGGCATCATCAGCTGGGGCCCTGGGCTGTGACAGAGGATGTCCCGGGTG	1258
Qy	1076	TGTACACAAAGGTTACCAACTTACCTAGACTGGATTTCGTGACCAATGCGACCG	1128
Db	1259	TGTACACAAAGGTTACCAACTTACCTAGACTGGATTTCGTGACCAATGCGACCG	1311
RESULT 3			
AAQ06762			
ID	AAQ06762 standard; DNA; 1341 BP.		
XX	AAQ06762;		
XX	AC		
XX	AC		
DT	09-JAN-2003 (updated)		
DT	01-MAR-1991 (first entry)		
XX	Sequence encoding non-glycosylated tPA deriv.		
XX	Tissue plasminogen activator; thrombolysis; finger domain;		
KW	EGF domain; ss.		
KW	EGF domain; ss.		
XX	Synthetic.		
XX	Key		
FH	Location/Qualifiers		
FT	4..1341		
FT	/*tag= a		
FT	/product=tPA deriv.		
XX	EP400545-A.		
XX	05-DEC-1990.		
XX	28-MAY-1990; 90EP-0110096.		
XX	14-JUL-1989; 89DE-3923339.		
PR	31-MAY-1989; 89DE-3917781.		
XX	(BOEF ) BOEHRINGER MANNHEIM GMBH.		
XX	Stern A, Kohnert U, Rudolph R, Fischer S, Martin U;		
XX	WPI; 1990-363094/49.		
DR	P-PSDB; AAR08150.		
XX	New non-glycosylated form of tissue plasminogen activator - with		
FT	thrombolytic activity and long plasma life		
XX	Claim 2; page 16; 21pp; German.		
XX	This sequence encodes a non-glycosylated tPA deriv. lacking the		
CC	finger- and EGF-domains. The polypeptide has a lower clearance		
CC	rate (longer half-life) than natural tPA while retaining thrombo-		
CC	lytic activity and stimulation by fibrin.		
CC	(updated on 09-JAN-2003 to add missing OS field.)		
XX	Sequence 1341 BP; 293 A; 397 C; 399 G; 252 T; 0 other;		
XX	Query Match		
XX	Best Local Similarity 94.4%; Score 1065; DB 11; Length 1341;		
XX	Matches 1068; Conservative		
XX	Indels 0; Mismatches 5; Gaps 0;		

56 CCCAGCGCGCTCTCAGGGAACAGTACTGCTACTTTGGGAATGGTTCAGCTACCGTG 115  
116 GCACGACAGCTCTACCGAGTGGGTGCTCTGCTCCCGTGGAAATTCATGATCTTGA 175  
329 GCACGACAGCTCTACCGAGTGGGTGCTCTGCTCCCGTGGAAATTCATGATCTTGA 388  
176 TAGGCAAGGTTTACACGACACAGAACCCGAGTGGCCAGGCACTGGCTGGGCAACATA 235  
389 TAGGCAAGGTTTACACGACACAGAACCCGAGTGGCCAGGCACTGGCTGGGCAACATA 448  
236 ATTACTGCGGAATCCTGATGGGATGCCAAGCCCTGTGTCACGCTGCTGAAGAACCGCA 295  
449 ATTACTGCGGAATCCTGATGGGATGCCAAGCCCTGTGTCACGCTGCTGAAGAACCGCA 508  
296 GGCTGACGTGGAGTACTGTGATGTCCTCTGCTCCACCTGCGGCTGAGACAGTACA 355  
509 GGCTGACGTGGAGTACTGTGATGTCCTCTGCTCCACCTGCGGCTGAGACAGTACA 568  
356 GCCAGCCTCAGTTTCGATCAAGAGAGGCTTTCGCGGACATCGCTCCACCCCTGGC 415  
569 GCCAGCCTCAGTTTCGATCAAGAGAGGCTTTCGCGGACATCGCTCCACCCCTGGC 628  
416 AGGCTGCCATCTTTGCCAAGCACAGAGTGGCCCGGAGCGGTTCTGTGGGGGCA 475  
629 AGGCTGCCATCTTTGCCAAGCACAGAGTGGCCCGGAGCGGTTCTGTGGGGGCA 688  
476 TACTCATCAGCTCTGCTGGATTCTCTCTGCGGCCACTGCTTCAGGAGGTTTCGCG 535  
689 TACTCATCAGCTCTGCTGGATTCTCTCTGCGGCCACTGCTTCAGGAGGTTTCGCG 748  
536 CCCACACCTGACGCTGATCTTGGGACAGACATACCGGGTGGTCCCTGGGAGGAGGC 595  
749 CCCACACCTGACGCTGATCTTGGGACAGACATACCGGGTGGTCCCTGGGAGGAGGC 808  
596 AGAAATTTGAAGTTCGAAATATATGTCATAGGAATTCGATGATGACACTTACGACA 655  
809 AGAATTTGAAGTTCGAAATATATGTCATAGGAATTCGATGATGACACTTACGACA 868  
656 ATGACATTCGCTGCTGAGTGAATTCGTCGCGGCCACTGCTTCAGGAGAGGTTTCGCG 715  
869 ATGACATTCGCTGCTGAGTGAATTCGTCGCGGCCACTGCTTCAGGAGAGGTTTCGCG 928  
716 TGCTCCGACCTGCTGCTGCTTCCCGCGGACCTGACCTGCGGACCTGAGAGGTTG 775  
929 TGCTCCGACCTGCTGCTGCTTCCCGCGGACCTGACCTGCGGACCTGAGAGGTTG 988  
776 AGCTCTCGGCTACGCGAAGATGAGGCTTGTCTCTCTTCTATTTCGAGCGGCTGAAG 835  
989 AGCTCTCGGCTACGCGAAGATGAGGCTTGTCTCTCTTCTATTTCGAGCGGCTGAAG 1048  
836 AGGCTCATGCTGAGTACCATCCAGCCGCTGCATCAACATTTACTTAAACAGAA 895  
1049 AGGCTCATGCTGAGTACCATCCAGCCGCTGCATCAACATTTACTTAAACAGAA 1108  
896 CAGTACACGACAAATGCTGTGTGTGAGACACTCGAGAGCGGGGCCCCCAGGCAACT 955  
1109 CAGTACACGACAAATGCTGTGTGTGAGACACTCGAGAGCGGGGCCCCCAGGCAACT 1168  
956 TGCAGAGCGCTTCCAGGCGGATTCGGAGGCGGCTGCTGTGTCTGAACGATGGCGCA 1015  
1169 TGCAGAGCGCTTCCAGGCGGATTCGGAGGCGGCTGCTGTGTCTGAACGATGGCGCA 1228  
1016 TGACTTTGGTGGGATCATCAGTGGGCTTGGGCTGTGGAAGAGATGTCGCGGTG 1075  
1229 TGACTTTGGTGGGATCATCAGTGGGCTTGGGCTGTGGAAGAGATGTCGCGGTG 1288  
1076 TGTACAAAGGTTTACCACTTACCTAGACTGGATTCGTGACACATGCGACCG 1128  
1289 TGTACAAAGGTTTACCACTTACCTAGACTGGATTCGTGACACATGCGACCG 1341

RESULT 4  
AAN91133  
ID AAN91133 standard; DNA; 1419 BP.  
XX  
AC AAN91133;  
XX  
DT 25-MAR-2003 (updated)  
DT 03-OCT-2002 (updated)  
DT 18-JUN-1990 (first entry)  
XX  
DE Sequence of coding region in plasmid pmTQk112.  
XX  
KW Tissue plasminogen activator; tPA; thrombolytic agent;  
KW plasminogen; vascular diseases.  
XX  
OS Synthetic.  
XX  
FH Key Location/Qualifiers  
FT 1..1065  
FT /\*tag= a  
XX  
PN EP302456-A.  
XX  
PD 08-FEB-1989.  
XX  
PE 02-AUG-1988; 88EP-0112569.  
XX  
PR 03-AUG-1987; 87GB-0018298.  
PR 26-OCT-1987; 87GB-0025052.  
PR 13-NOV-1987; 87GB-0026683.  
XX  
PA (FUJI ) FUJISAWA PHARM CO LTD.  
XX  
PI Niwa M, Saito Y, Sasaki H, Hayashi M, Notani J, Kobayashi M;  
XX  
DR WPI; 1989-040625/06.  
XX  
P-PSDB; AAP94416.  
XX  
PT New tissue plasminogen activator -  
PT comprising finger and growth factor domains lacking tPA for  
PT longer half-life and stronger thrombolytic activity.  
XX  
PS Disclosure; Page ?; 68pp; English.  
XX  
CC (Updated on 03-OCT-2002 to add missing OS field.)  
CC (Updated on 25-MAR-2003 to correct PA field.)  
XX  
SQ Sequence 1419 BP; 309 A; 413 C; 426 G; 271 T; 0 other;  
Query Match 94.4%; Score 1065; DB 10; Length 1419;  
Best Local Similarity 99.5%; Pred. No. 3.9e-225;  
Matches 1068; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
QY 56 CCCAGCGCGCTCTCAGGGAACAGTACTGCTACTTTGGGAATGGTTCAGCTACCGTG 115  
DB 344 CCCCTGCTGCTCTGAGGGAACAGTACTGCTACTTTGGGAATGGTTCAGCTACCGTG 403  
QY 116 GCACGACAGCTCTACCGAGTGGGTGCTCTGCTCCCGTGGAAATTCATGATCTTGA 175  
DB 404 GCACGACAGCTCTACCGAGTGGGTGCTCTGCTCCCGTGGAAATTCATGATCTTGA 463  
QY 176 TAGGCAAGGTTTACACGACACAGAACCCGAGTGGCCAGGCACTGGCTGGGCAACATA 235  
DB 464 TAGGCAAGGTTTACACGACACAGAACCCGAGTGGCCAGGCACTGGCTGGGCAACATA 523  
QY 236 ATTACTGCGGAATCCTGATGGGATGCCAAGCCCTGTGTCACGCTGCTGAAGAACCGCA 295  
DB 524 ATTACTGCGGAATCCTGATGGGATGCCAAGCCCTGTGTCACGCTGCTGAAGAACCGCA 583  
QY 296 GGCTGACGTGGAGTACTGTGATGTCCTCTGCTCCACCTGCGGCTGAGACAGTACA 355  
DB 584 GGCTGACGTGGAGTACTGTGATGTCCTCTGCTCCACCTGCGGCTGAGACAGTACA 643

QY 356 GCCAGCCTCAGTTTCGATCAAGAGAGGGCTCTTTCGCCGACATCGCTCCCAACCCCTGGC 415  
 Db 644 GCCAGCCTCAGTTTCGATCAAGAGAGGGCTCTTTCGCCGACATCGCTCCCAACCCCTGGC 703  
 QY 416 AGGCTGCCATCTTTGCCAAGCACAGAGAGTCCCGGAGAGGGTTCCTGTGCGGGGCA 475  
 Db 704 AGGCTGCCATCTTTGCCAAGCACAGAGAGTCCCGGAGAGGGTTCCTGTGCGGGGCA 763  
 QY 476 TACTCATCAGCTCTCTGCTGGATTTCTCTGCGCGCCCACTGCTTCCAGGAGAGGTTTCGCG 535  
 Db 764 TACTCATCAGCTCTCTGCTGGATTTCTCTGCGCGCCCACTGCTTCCAGGAGAGGTTTCGCG 823  
 QY 536 CCCACCCTCAGCGTGATCTTTGGGCGAACAATACCGGGTGCTCCCTGGGAGAGGAGC 595  
 Db 824 CCCACCCTCAGCGTGATCTTTGGGCGAACAATACCGGGTGCTCCCTGGGAGAGGAGC 883  
 QY 596 AGAATTTGAAGTCCAAAGATACATGTCCTATAGGAATTCGATGACACTTACGACA 655  
 Db 884 AGAATTTGAAGTCCAAAGATACATGTCCTATAGGAATTCGATGACACTTACGACA 943  
 QY 656 ATGACATTCGCTGCTGACGCTGAAATCGGATTCGCTCCGCTGTGCCAGGAGCAGCG 715  
 Db 944 ATGACATTCGCTGCTGACGCTGAAATCGGATTCGCTCCGCTGTGCCAGGAGCAGCG 1003  
 QY 716 TGGTCCGCACTGTGTGCTTCCCGCGGACCTGCAGCTGCGGACTGGACGAGTG 775  
 Db 1004 TGGTCCGCACTGTGTGCTTCCCGCGGACCTGCAGCTGCGGACTGGACGAGTG 1063  
 QY 776 AGCTTCGCGCTACGCGAGCATGAGGCTTGTCTCTTCTATTCGAGCGGCTGAAGG 835  
 Db 1064 AGCTTCGCGCTACGCGAGCATGAGGCTTGTCTCTTCTATTCGAGCGGCTGAAGG 1123  
 QY 836 AGGCTCATGTGACAGTGTACCATCCAGCGCTGCACATCAACATTTACTTAACAGAA 895  
 Db 1124 AGGCTCATGTGACAGTGTACCATCCAGCGCTGCACATCAACATTTACTTAACAGAA 1183  
 QY 896 CAGTCACCGCAACATGTGTGTGTGGAGACATTCGAGCGGGGGCCCCCAGGCAAACT 955  
 Db 1184 CAGTCACCGCAACATGTGTGTGTGGAGACATTCGAGCGGGGGCCCCCAGGCAAACT 1243  
 QY 956 TGCAGAGCGCTGCCAGGCGGATTCGGAGAGCGCCCTGCTGTGTGACAGATGCGCGCA 1015  
 Db 1244 TGCAGAGCGCTGCCAGGCGGATTCGGAGAGCGCCCTGCTGTGTGACAGATGCGCGCA 1303  
 QY 1016 TGACTTTGGTGGGCATCATCAGCTGGGCGCTGTGGACAGAGGATGTCCCGGCTG 1075  
 Db 1304 TGACTTTGGTGGGCATCATCAGCTGGGCGCTGTGGACAGAGGATGTCCCGGCTG 1363  
 QY 1076 TGTACAAAGGTTACCAACTACCTAGACTGGATTCGTGACAAATGCGACCG 1128  
 Db 1364 TGTACAAAGGTTACCAACTACCTAGACTGGATTCGTGACAAATGCGACCG 1416

## RESULT 5

AA82178  
 ID AA82178 standard; DNA; 1689 BP.

XX AC AA82178;

XX DT 25-MAR-2003 (updated)

XX DT 10-MAR-2003 (updated)

XX DT 03-NOV-1990 (first entry)

XX DE Tissue plasminogen activator with S-119 substd for M.

XX KW Tissue plasminogen activator; tPA; thrombosis; N-glycosylation; ss..

XX OS Unidentified.

XX FH Key Location/Qualifiers

XX FT CDS 1..1689

XX FT /\*tag= a

/product=modified tPA

JP63230083-A.

26-SEP-1988.

20-MAR-1987; 87JP-0064339.

20-MAR-1987; 87JP-0064339.

(EISA ) EISAI CO LTD.

WPI; 1988-311961/44.

P-PSDB; AAP82581.

Modified tissue plasminogen activator - having glycine-183 and serine-186 residues substd. with serine and threonine.

Disclosure; Page ?; 16pp; Japanese.

One N-glycosylation site, i.e. NSS (117-119) is substituted with NSM and the N-glycosylation is removed.

Plasmid encoding the modified tPA is 99-6300 and its transformant is E.coli RRI-Zem 99-6300 (FERM P-9127).

This modified tPA, used to treat thrombosis, is of high quality and has a longer half life period in blood.

See also AAN82177-N82179.

(Updated on 10-MAR-2003 to add missing OS field.)

(Updated on 25-MAR-2003 to correct PF field.)

(Updated on 25-MAR-2003 to correct PR field.)

(Updated on 25-MAR-2003 to correct PA field.)

Sequence 1689 BP; 376 A; 482 C; 504 G; 327 T; 0 other;

Query Match 94.4%; Score 1065; DB 9; Length 1689;

Best Local Similarity 95.5%; Pred. No. 4e-225;

Matches 1068; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 56 CCCAGGCGGCTCTGAGGAAACAGTACTGCTACTTTGGAAATGGGTACGCTACCGTG 115

Db 614 CCCCTGCTCTCTGAGGAAACAGTACTGCTACTTTGGAAATGGGTACGCTACCGTG 673

QY 116 GCACGCACAGCTCACCAGTCTGCTCTCCCTCCCTGGAATTCATGATCTCTGA 175

Db 674 GCACGCACAGCTCACCAGTCTGCTCTCCCTCCCTGGAATTCATGATCTCTGA 733

QY 176 TAGCAGAGTTTACACAGCACAGAACCCAGTCCGAGGACTGGGCTGGGCAACATA 235

Db 734 TAGCAGAGTTTACACAGCACAGAACCCAGTCCGAGGACTGGGCTGGGCAACATA 793

QY 236 ATTAATGCGGGAATCTCTGATGGGATGCAAGCCCTGGTGCACGTCTGAAGAACCGCA 295

Db 794 ATTAATGCGGGAATCTCTGATGGGATGCAAGCCCTGGTGCACGTCTGAAGAACCGCA 853

QY 296 GGCTGACGTGGAGTACTGTGATGTCCTCTCTCCCTCCCTGCGGCTGAGACAGTACA 355

Db 854 GGCTGACGTGGAGTACTGTGATGTCCTCTCTCCCTCCCTGCGGCTGAGACAGTACA 913

QY 356 GCCAGCCTCAGTTTCGATCAAGAGAGGGCTCTTTCGCCGACATCGCTCCCAACCCCTGGC 415

Db 914 GCCAGCCTCAGTTTCGATCAAGAGAGGGCTCTTTCGCCGACATCGCTCCCAACCCCTGGC 973

QY 416 AGGCTGCCATCTTTGCCAAGCACAGAGAGTCCCGGAGAGCGGTTCCTGTGCGGGGCA 475

Db 974 AGGCTGCCATCTTTGCCAAGCACAGAGAGTCCCGGAGAGCGGTTCCTGTGCGGGGCA 1033

QY 476 TACTCATCAGCTCTCTGCTGGATTTCTCTGCGGCCCACTGCTTCCAGAGAGGTTTCGCG 535

Db 1034 TACTCATCAGCTCTCTGCTGGATTTCTCTGCGGCCCACTGCTTCCAGAGAGGTTTCGCG 1093

QY 536 CCCACCACTGACGCTGATCTTGGGCGAACAATACCGGGTGCTCCCTGGGAGAGGAGC 595

Db 1094 CCCACCACTGACGCTGATCTTGGGCGAACAATACCGGGTGCTCCCTGGGAGAGGAGC 1153

QY 596 AGAAATTTGAAGTCGAAATAATATTGTCATAGGAATTCGATGATGACACTTACGACA 655  
 DB 1154 AGAAATTTGAAGTCGAAATAATATTGTCATAGGAATTCGATGATGACACTTACGACA 1213  
 QY 656 ATGACATTCGGCTGCTGACGCTGAAATCGATTCCTCCGCTGTCGCCAGAGAGACGCG 715  
 DB 1214 ATGACATTCGGCTGCTGACGCTGAAATCGATTCCTCCGCTGTCGCCAGAGAGACGCG 1273  
 QY 716 TGGTCGGCACTGTGTCCTTCCCGCGGACCTGACGCTGCGGACCTGACGAGTGTG 775  
 DB 1274 TGGTCGGCACTGTGTCCTTCCCGCGGACCTGACGCTGCGGACCTGACGAGTGTG 1333  
 QY 776 AGCTCTCCGGCTACCGCAAGCATGAGGCTTGTCTCTTCTTATTCGAGCGGCTGAAGG 835  
 DB 1334 AGCTCTCCGGCTACCGCAAGCATGAGGCTTGTCTCTTCTTATTCGAGCGGCTGAAGG 1393  
 QY 836 AGGCTCATGTGACATGTACCATCCAGCCGCTGCACATCAACAATTTACTTTAACAGAA 895  
 DB 1394 AGGCTCATGTGACATGTACCATCCAGCCGCTGCACATCAACAATTTACTTTAACAGAA 1453  
 QY 896 CAGTCACCGACACATGCTGTGTGTCGAGACACTCGGAGCGCGGCCCGCCAGGCAAACT 955  
 DB 1454 CAGTCACCGACACATGCTGTGTGTCGAGACACTCGGAGCGCGGCCCGCCAGGCAAACT 1513  
 QY 956 TGCACGACGCTCCAGGCGGATTCGGAGGCGCCCTGTGTGTCTGAACGATGCGCGCA 1015  
 DB 1514 TGCACGACGCTCCAGGCGGATTCGGAGGCGCCCTGTGTGTCTGAACGATGCGCGCA 1573  
 QY 1016 TGACTTTGGTGGGCAATCATGCTAGCTGGGCGCTGGGCTGTGGACAGAGATGTCGCGGTG 1075  
 DB 1574 TGACTTTGGTGGGCAATCATGCTAGCTGGGCGCTGGGCTGTGGACAGAGATGTCGCGGTG 1633  
 QY 1076 TGTACAAAGGTTACCACTACCTAGACTGATTCGTCGACACATGCGACCG 1128  
 DB 1634 TGTACAAAGGTTACCACTACCTAGACTGATTCGTCGACACATGCGACCG 1686  
 RESULT 6  
 ID AAQ01358  
 XX AAQ01358 standard; DNA; 1780 BP.  
 AC AAQ01358;  
 XX  
 XX 25-MAR-2003 (updated)  
 DT 08-SEP-1990 (first entry)  
 XX  
 DE Sequence encoding wild type tissue plasminogen activator (t-PA).  
 XX  
 KW Wild type tissue plasminogen activator (t-PA); infarction treatment;  
 KW thrombosis treatment; embolism treatment.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT misc\_feature 1..123  
 FT /tag= a  
 FT /note="synthetic 5' adaptor"  
 FT misc\_feature 1738..1780  
 FT /tag= b  
 FT /note="synthetic 3' adaptor"  
 XX  
 PN EP351246-A.  
 XX  
 PD 17-JAN-1990.  
 XX  
 PF 14-JUL-1989; 89EP-0307194.  
 XX  
 PR 15-JUL-1988; 88DK-0003952.  
 XX  
 PA (NOVO ) NOVO-NORDISK AS.  
 XX  
 PX Petersen LC, Boel E;

XX WPI; 1990-016567/03.  
 DR P-PSDB; AAR04699.  
 XX  
 PT New tissue plasminogen activator (t-PA) analogue -  
 PT with higher fibrin selectivity than native t-PA, useful for  
 PT treating infarction, thrombosis and embolism  
 XX  
 PS Disclosure; Fig 5A-D; 24pp; English.  
 XX  
 CC It is modified in the patent by replacing one or more codons specifying  
 CC a positively charged amino acid. The resulting analogues are inserted  
 CC into a replicable expression vector which is used to transform or  
 CC transfect a host cell which is grown to express a t-PA analogue. The  
 CC t-PA analogue is useful for the treatment of diseases or disorders  
 CC associated with the formation of thrombi in blood vessels e.g.  
 CC infarctions, thrombosis and embolism. The analogue, in the 1-chain form,  
 CC exhibits the properties of a proenzyme; however on plasmin-catalysed  
 CC cleavage of the one-chain form, the activity of the 2-chain form is fully  
 CC retained. Compared to native t-PA, it has a higher fibrin selectivity,  
 CC this results in a fibrinolytic agent with a higher fibrin selectivity,  
 CC as the fibrinogenolytic activity induced by the analogue is reduced  
 CC relative to that induced by native 1-chain t-PA.  
 CC (Updated on 25-MAR-2003 to correct PA field.)  
 XX  
 SQ Sequence 1780 BP; 405 A; 513 C; 523 G; 339 T; 0 other;  
 Query March 94.4%; Score 1065; DB 11; Length 1780;  
 Best Local Similarity 99.5%; Pred. No. 4e-225;  
 Matches 1068; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
 QY 56 CCCAGGCGGCTCTGAGGGAACAGTCACTGCTACTTTGGGAATGGGTGACGCTTACCGTG 115  
 DB 630 CCCCTGCTGCTCTGAGGGAACAGTCACTGCTACTTTGGGAATGGGTGACGCTTACCGTG 689  
 QY 116 GCACGCAAGCCTCACCGAGTCGGGTGCTCTCTGCTCCCGTGGAAATTCATGATCTCTGA 175  
 DB 690 GCACGCAAGCCTCACCGAGTCGGGTGCTCTCTGCTCCCGTGGAAATTCATGATCTCTGA 749  
 QY 176 TAGCAGAGTTTACACAGGCAAGAACCCAGTGCAGGCACTGGGCTGGGCAACATA 235  
 DB 750 TAGCAGAGTTTACACAGGCAAGAACCCAGTGCAGGCACTGGGCTGGGCAACATA 809  
 QY 236 ATTAAGTCGCGGAATCTCTGATGGGATGCAAGCCCTGCTGCAAGAACCGCA 295  
 DB 810 ATTAAGTCGCGGAATCTCTGATGGGATGCAAGCCCTGCTGCAAGAACCGCA 869  
 QY 296 GGCTGACGTGGAGTACTGTGATGTGCTCTCTGCTCCAGCTGCGGCTGAGACAGTACA 355  
 DB 870 GGCTGACGTGGAGTACTGTGATGTGCTCTCTGCTCCAGCTGCGGCTGAGACAGTACA 929  
 QY 356 GCCAGCCTCAGTTTCGCATCAAGAGGAGGCTCTTCGCGGACATCGCTCCACCCCTGGC 415  
 DB 930 GCCAGCCTCAGTTTCGCATCAAGAGGAGGCTCTTCGCGGACATCGCTCCACCCCTGGC 989  
 QY 416 AGGCTGCATCTTTGCGCAAGCACAGGAGGTGCGCCGAGAGCGGTTCCTGTCGCGGGGCA 475  
 DB 990 AGGCTGCATCTTTGCGCAAGCACAGGAGGTGCGCCGAGAGCGGTTCCTGTCGCGGGGCA 1049  
 QY 476 TACTCATCAGCTCCTGCTGGAATTCCTCTGCGGCCACTGCTTCCAGAGAGGTTTCGCG 535  
 DB 1050 TACTCATCAGCTCCTGCTGGAATTCCTCTGCGGCCACTGCTTCCAGAGAGGTTTCGCG 1109  
 QY 536 CCCACCACTGACGGTGTATCTTGGGCAAGACATACCGGTGCTCCCTGGCGAGGAGGAGC 595  
 DB 1110 CCCACCACTGACGGTGTATCTTGGGCAAGACATACCGGTGCTCCCTGGCGAGGAGGAGC 1169  
 QY 596 AGAAATTTGAAGTCGAAATAATATTGTCATAGGAATTCGATGATGACACTTACGACA 655  
 DB 1170 AGAAATTTGAAGTCGAAATAATATTGTCATAGGAATTCGATGATGACACTTACGACA 1229  
 QY 656 ATGACATTCGGCTGCTGACGCTGAAATCGATTCCTCCGCTGTCGCCAGAGAGGAGCG 715

Db 1230 ATGACATTGCGCTGCTGACGTGAATCGGATTCGCTCCGCTGTGCCCGCAGGAGCAGCG 1289  
 QY 716 TGGTCCGCACTGTGTCCTTCCCGCGGAGACTCGAGCTGCCGAGCTGGAGGAGTGTG 775  
 Db 1290 TGGTCCGCACTGTGTCCTTCCCGCGGAGACTCGAGCTGCCGAGCTGGAGGAGTGTG 1349  
 QY 776 AGCTCTCCGCTACCGCAAGCATGAGGCTTGTCTCTTTCTATTTCGGAGCGGCTGAAGG 835  
 Db 1350 AGCTCTCCGCTACCGCAAGCATGAGGCTTGTCTCTTTCTATTTCGGAGCGGCTGAAGG 1409  
 QY 836 AGGCTCATGTGACAGTGTACCCATCCAGCGCTGCACATCAACAATTTACTTTAACAGAA 895  
 Db 1410 AGGCTCATGTGACAGTGTACCCATCCAGCGCTGCACATCAACAATTTACTTTAACAGAA 1469  
 QY 896 CAGTCACCGCAACATGTGTGTGTGGAGACATCGGAGCGCGGCGCCCGCAAACT 955  
 Db 1470 CAGTCACCGCAACATGTGTGTGTGGAGACATCGGAGCGCGGCGCCCGCAAACT 1529  
 QY 956 TGCACGACGCTGCCAGGCGGATTCGGAGGCGCCCTGTGTGTGTGAACGATGCGCGCA 1015  
 Db 1530 TGCACGACGCTGCCAGGCGGATTCGGAGGCGCCCTGTGTGTGTGAACGATGCGCGCA 1589  
 QY 1016 TGACTTTGGTGGGCATCATCAGCTCGGCGCTGGGCTGTGGACAGAGGATGTCCCGGCTG 1075  
 Db 1590 TGACTTTGGTGGGCATCATCAGCTCGGCGCTGGGCTGTGGACAGAGGATGTCCCGGCTG 1649  
 QY 1076 TGTACACAAAGTTTACCACTACCTAGACTGGATTTCGTGACAAATGCGACCG 1128  
 Db 1650 TGTACACAAAGTTTACCACTACCTAGACTGGATTTCGTGACAAATGCGACCG 1702

## RESULT 7

ID AAN91119 standard; DNA; 2100 BP.  
 AC AAN91119;

DT 25-MAR-2003 (updated)  
 DT 03-OCT-2002 (updated)  
 DT 18-JUN-1990 (first entry)

Sequence of native tPA in plasmid pST112.

Tissue plasminogen activator; tPA; thrombolytic agent;  
 plasminogen; vascular diseases.

Synthetic.

Key Location/Qualifiers  
 CDS 25..1710  
 FT /\*tag= a

EP302456-A.

08-FEB-1989.

02-AUG-1988; 88EP-0112569.

03-AUG-1987; 87GB-0018298.

26-OCT-1987; 87GB-0025052.

13-NOV-1987; 87GB-0026683.

(FUJII) FUJISAWA PHARM CO LTD.

NIWA M, Saito Y, Sasaki H, Hayashi M, Notani J, Kobayashi M;

WPI; 1989-040625/06.

P-PSDB; AAP94406.

New tissue plasminogen activator -

comprising finger and growth factor domains lacking tPA for  
 longer half-life and stronger thrombolytic activity.

XX

PS Example 29; Fig 21; 68pp; English.

XX cDNA sequence of native tPA gene is excised from plasmid pST112, and  
 CC digested with BamHI and SalI to form plasmid pST118.  
 CC (Updated on 03-OCT-2002 to add missing OS field.)  
 XX (Updated on 25-MAR-2003 to correct PA field.)

SQ Sequence 2100 BP; 498 A; 595 C; 597 G; 410 T; 0 other;

Query Match 94.4%; Score 1065; DB 10; Length 2100;  
 Best Local Similarity 99.5%; Pred. No. 4,1e-225;  
 Matches 1068; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 56 CCCAGGCGCTCTGAGGAAACAGTACTCTACTTTGGGAATGGGTCAACCTACCGTG 115  
 Db 638 CCCCTGCGCTCTGAGGAAACAGTACTCTACTTTGGGAATGGGTCAACCTACCGTG 697  
 QY 116 GCAGCAGAGCTCACCGAGTGGGTGCTCTGCTCCCTGGAATTCATGATCTCTGA 175  
 Db 698 GCAGCAGAGCTCACCGAGTGGGTGCTCTGCTCCCTGGAATTCATGATCTCTGA 757  
 QY 176 TAGCAAGGTTTACACAGCAGAACCCAGTCCCGAGGCACTGGGCTGGGCAACATA 235  
 Db 758 TAGCAAGGTTTACACAGCAGAACCCAGTCCCGAGGCACTGGGCTGGGCAACATA 817  
 QY 236 ATTACTGCGGAATCTGATGGGATGCCAAGCCCTGGTGCACGTCTGAAGAACCGCA 295  
 Db 818 ATTACTGCGGAATCTGATGGGATGCCAAGCCCTGGTGCACGTCTGAAGAACCGCA 877  
 QY 296 GGCCTGAGCTGGGAGTACTGTGATGTGCTCTGCTCCCTGCTCCAGTGGGCTGAGACAGTACA 355  
 Db 878 GGCCTGAGCTGGGAGTACTGTGATGTGCTCTGCTCCCTGCTCCAGTGGGCTGAGACAGTACA 937  
 QY 356 GCCAGCCTCAGTTTCGCATCAAGAGGAGGCTCTTCGCGAGCATCGCCTCCACCCCTGGC 415  
 Db 938 GCCAGCCTCAGTTTCGCATCAAGAGGAGGCTCTTCGCGAGCATCGCCTCCACCCCTGGC 997  
 QY 416 AGGCTGCCATCTTTGCCAAGCAGAGGAGTCCGCCGAGAGCGGTTCCTGTGCGGGGCA 475  
 Db 998 AGGCTGCCATCTTTGCCAAGCAGAGGAGTCCGCCGAGAGCGGTTCCTGTGCGGGGCA 1057  
 QY 476 TACTCATCAGCTCTGTGTGATTCCTGTGCGGCCCACTGTCTCAGAGAGGTTTCGCG 535  
 Db 1058 TACTCATCAGCTCTGTGTGATTCCTGTGCGGCCCACTGTCTCAGAGAGGTTTCGCG 1117  
 QY 536 CCCACCACTGACGCTGATCTTGGGAGAACATACCGGCTGCTCCCTGGCGAGGAGGAGC 595  
 Db 1118 CCCACCACTGACGCTGATCTTGGGAGAACATACCGGCTGCTCCCTGGCGAGGAGGAGC 1177  
 QY 596 AGAAATTTGAATCGAATAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA 655  
 Db 1178 AGAAATTTGAATCGAATAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA 1237  
 QY 656 ATGACATGCGCTCTGCGAGCTGAAATCGGAATCGTCCCGCTGTGCCAGAGAGCAGCG 715  
 Db 1238 ATGACATGCGCTCTGCGAGCTGAAATCGGAATCGTCCCGCTGTGCCAGAGAGCAGCG 1297  
 QY 716 TGGTCCGCACTGTGTGCTTCCCGGGGAGACCTGCAGCTCCGAGCTGGAGCGAGTGTG 775  
 Db 1298 TGGTCCGCACTGTGTGCTTCCCGGGGAGACCTGCAGCTCCGAGCTGGAGCGAGTGTG 1357  
 QY 776 AGCTCTCCGCTACGCAAGCATGAGGCTTGTCTCTTTCTATTTCGAGCGGCTGAAGG 835  
 Db 1358 AGCTCTCCGCTACGCAAGCATGAGGCTTGTCTCTTTCTATTTCGAGCGGCTGAAGG 1417  
 QY 836 AGGCTCATGTGACAGTGTACCCATCCAGCGCTGCACATCAACAATTTACTTTACAGAA 895  
 Db 1418 AGGCTCATGTGACAGTGTACCCATCCAGCGCTGCACATCAACAATTTACTTTACAGAA 1477  
 QY 896 CAGTCACCGCAACATGCTGTGTGCTGGAGACATCGGAGCGGCGGCGCCCGCAAACT 955  
 Db 1478 CAGTCACCGCAACATGCTGTGTGCTGGAGACATCGGAGCGGCGGCGCCCGCAAACT 1537

QY 956 TGCACGAGCCTGCCAGGCGGATTCGGAGGCCCTCGTGTGTCTGAACGATGCGCCGCA 1015  
 Db 1538 TGCACGAGCCTGCCAGGCGGATTCGGAGGCCCTCGTGTGTCTGAACGATGCGCCGCA 1597  
 QY 1016 TGACTTTGGTGGGATCATCAGCTGGGCGCTGGGCTGTGGACAGAGATGTCCGGGTG 1075  
 Db 1598 TGACTTTGGTGGGATCATCAGCTGGGCGCTGGGCTGTGGACAGAGATGTCCGGGTG 1657  
 QY 1076 TGTACACAAAGGTTACCAACTACCTAGACTGGATTTCGTGACAAATGCGACCG 1128  
 Db 1658 TGTACACAAAGGTTACCAACTACCTAGACTGGATTTCGTGACAAATGCGACCG 1710

## RESULT 8

AAQ05532  
 ID AAQ05532 standard; DNA; 2100 BP.

XX AC AAQ05532;

XX DT 25-MAR-2003 (updated)  
 XX DT 11-DEC-1990 (first entry)

XX DE Plasmid pST112 encoding novel N-terminal for tissue plasminogen  
 XX DE activator (tPA).

XX KW Fibrin; clotting; thrombolytic; vascular disease; stroke;  
 XX KW myocardial infarction; heart attack; pulmonary embolism; ds;

XX OS Homo sapiens.

XX FH Key Location/Qualifiers  
 XX FT CDS 25..1710  
 XX FT mat\_peptide /\*tag= a  
 FT 130..1710  
 FT /\*tag= b

XX FN EP379890-A.

XX PD 01-AUG-1990.

XX PF 10-JAN-1990; 90EP-0100457.  
 XX PR 23-JAN-1989; 89GB-0001422.

XX PA (FUJI ) FUJISAWA PHARM CO LTD.  
 XX PI Niwa M, Satoh S, Suzuki S, Otsuka K, Kusunoki C;

XX DR WPI; 1990-232757/31.

XX DR P-PSDB; AAR06237.

XX PT New tissue plasminogen activator - having N-terminal peptide of  
 PT plasminogen linked to tissue plasminogen activator for increased  
 PT stability in vivo.

XX PS Disclosure; Page ?; ?pp; English.

XX CC tPA with N-terminal peptide of plasminogen linked is more stable in  
 CC vivo than the native form. It is useful as a thrombolytic agent in  
 CC the treatment of vascular diseases eg myocardial infarction,  
 CC pulmonary embolism etc.  
 CC (Updated on 25-MAR-2003 to correct PA field.)

XX SQ Sequence 2100 BP; 498 A; 595 C; 597 G; 410 T; 0 other;

Query Match 94.4%; Score 1065; DB 11; Length 2100;  
 Best Local Similarity 99.5%; Pred. No. 4.1e-225;  
 Matches 1068; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 56 CCCAGCGGCTCTGAGGAAACAGTGACTGCTACTTTGGGAATGGGTACGCTACCGTG 115  
 Db 638 CCCCTGCTCTGAGGAAACAGTGACTGCTACTTTGGGAATGGGTACGCTACCGTG 697

## RESULT 9

AAQ87370

ID AAQ87370 standard; DNA; 2162 BP.

XX XX

QY 116 GCACGACAGCCTCACCGAGTCGGGTGCTCTCGCTCCCTCCGTGAAATTCATGATCTCTGA 175  
 Db 698 GCACGACAGCCTCACCGAGTCGGGTGCTCTCGCTCCCTCCGTGAAATTCATGATCTCTGA 757  
 QY 176 TAGCAAGGTTTACAGACACAGAAACCCAGTGCCTCCAGGCACTGGGCTGGGCAAAACATA 235  
 Db 758 TAGCAAGGTTTACAGACACAGAAACCCAGTGCCTCCAGGCACTGGGCTGGGCAAAACATA 817  
 QY 236 ATTACTCCCGGAATCCTGATGGGGATGCCAAGCCCTGGTGCACGTCGTGAAGAACCCGCA 295  
 Db 818 ATTACTCCCGGAATCCTGATGGGGATGCCAAGCCCTGGTGCACGTCGTGAAGAACCCGCA 877  
 QY 296 GGCTGACGTGGGAGTACTGTGATGTGCCCTCTCTCCACCTGCGGCTCGAGACAGTACA 355  
 Db 878 GGCTGACGTGGGAGTACTGTGATGTGCCCTCTCTCCACCTGCGGCTCGAGACAGTACA 937  
 QY 356 GCCAGCCTCAGTTTCGCATCAAGAGAGGGTCTTTGCGCGACATCGCTCCCACTCCGCG 415  
 Db 938 GCCAGCCTCAGTTTCGCATCAAGAGAGGGTCTTTGCGCGACATCGCTCCCACTCCGCG 997  
 QY 416 AGGCTGCCATCTTTGCCAAGCAGAGAGGTGCGCCGAGAGCGGTTCCTGTGCGGGGCA 475  
 Db 998 AGGCTGCCATCTTTGCCAAGCAGAGAGGTGCGCCGAGAGCGGTTCCTGTGCGGGGCA 1057  
 QY 476 TACTCATCAGCTCCTGTGGATTCTCTGCGCGCCACTGCTTCAGAGAGGTTTCGCG 535  
 Db 1058 TACTCATCAGCTCCTGTGGATTCTCTGCGCGCCACTGCTTCAGAGAGGTTTCGCG 1117  
 QY 536 CCCACCACTGACCGTGATCTTTGGGCAAGACATACCGGGTGGTCCCTGGCGAGGAGAGC 595  
 Db 1118 CCCACCACTGACCGTGATCTTTGGGCAAGACATACCGGGTGGTCCCTGGCGAGGAGAGC 1177  
 QY 596 AGAATTTGAAGTCGAAATACATCTCCATAGGAATTCGATGATGACACTTTACGACA 655  
 Db 1178 AGAATTTGAAGTCGAAATACATCTCCATAGGAATTCGATGATGACACTTTACGACA 1237  
 QY 656 ATGACATTTGGCTGCTGCAGCTGAAATCGGATTCTCCCGTGTGCCAGAGAGCAGCG 715  
 Db 1238 ATGACATTTGGCTGCTGCAGCTGAAATCGGATTCTCCCGTGTGCCAGAGAGCAGCG 1297  
 QY 716 TGGTCCGACTGTGTGCTTCCCGCGGACCTGACAGTGCCTGGACCTGGACGAGTGTG 775  
 Db 1298 TGGTCCGACTGTGTGCTTCCCGCGGACCTGACAGTGCCTGGACCTGGACGAGTGTG 1357  
 QY 776 AGCTCTCGGCTACGGCAAGCATGAGCCCTGTCTCTCTTCTATTTCGGAGCGGCTGAAGG 835  
 Db 1358 AGCTCTCGGCTACGGCAAGCATGAGCCCTGTCTCTTCTATTTCGGAGCGGCTGAAGG 1417  
 QY 836 AGGCTCATGTGACAGTGTACCCATCCAGCCGCTGCACATCACAACATTTACTTAACAGAA 895  
 Db 1418 AGGCTCATGTGACAGTGTACCCATCCAGCCGCTGCACATCACAACATTTACTTAACAGAA 1477  
 QY 896 CAGTCAACGACAAATCTGTGTGCTGGAGACATCTCGAGCGGCGGCGCCAGGCAAACT 955  
 Db 1478 CAGTCAACGACAAATCTGTGTGCTGGAGACATCTCGAGCGGCGGCGCCAGGCAAACT 1537  
 QY 956 TGACAGAGCGCTGCCAGGCGATTTCGGAGGCCCTCGTGTGTCTGAACCATGCGCCGCA 1015  
 Db 1538 TGACAGAGCGCTGCCAGGCGATTTCGGAGGCCCTCGTGTGTGTGAACATGCGCCGCA 1597  
 QY 1016 TGACTTTGGTGGGATCATCAGCTGGGCGCTGGGCTGTGGACAGAGATGTCCCGGGTG 1075  
 Db 1598 TGACTTTGGTGGGATCATCAGCTGGGCGCTGGGCTGTGGACAGAGATGTCCCGGGTG 1657  
 QY 1076 TGTACACAAAGGTTACCAACTACCTAGACTGGATTTCGTGACAAATGCGACCG 1128  
 Db 1658 TGTACACAAAGGTTACCAACTACCTAGACTGGATTTCGTGACAAATGCGACCG 1710

AC AAQ87370;  
 XX 25-MAR-2003 (updated)  
 DT 19-SEP-1995 (first entry)  
 XX Human tPA clone.  
 DE  
 XX Intron; recombination; combinatorial gene; trans-splicing;  
 KW gene therapy; polymerase chain reaction; PCR; primer; amplification;  
 KW tissue plasminogen activator; tPA; plasmid tPA-KS+; thrombolytic;  
 KW ds.  
 XX  
 OS Synthetic.  
 XX  
 XX Key Location/Qualifiers  
 FH sig\_peptide 82..334  
 FT /tag= a  
 FT /note= "signal sequence and finger-like domain"  
 FT CDS 335..447  
 FT /tag= b  
 FT /product= EGF-like domain  
 FT CDS 448..714  
 FT /tag= c  
 FT /product= Kringle-1 domain  
 FT CDS 715..972  
 FT /tag= d  
 FT /product= Kringle-2 domain  
 FT CDS 973..2162  
 FT /tag= e  
 FT /product= catalytic domain  
 XX  
 PN WO9507351-A1.  
 XX  
 XX 16-MAR-1995.  
 PD  
 XX 12-SEP-1994; 94WO-US10146.  
 PF  
 XX 10-SEP-1993; 93US-0119512.  
 PR  
 XX (HARD ) HARVARD COLLEGE.  
 PA  
 XX Jarrell KA;  
 XX WPI; 1995-123425/16.  
 DR  
 XX  
 XX New intron-mediated recombinant techniques - used for the  
 PT generation and selection of novel genes and gene prods. for use  
 PT in therapy  
 PT  
 XX  
 PS Example 4; Page 59-60; 87pp; English.  
 XX  
 XX A cDNA clone of human tissue plasminogen activator (tPA) was  
 CC amplified by PCR using the primers given in AAQ87368-69. The  
 CC amplified tPA DNA (AAQ87370) was ligated into vector KS+ to  
 CC obtain plasmid tPA-KS+. The construct was used in combinatorial  
 CC methods involving RNA splicing-mediated shuffling of tPA domains.  
 CC in plasmid PINVI (AAQ87347) to generate novel tPAs having  
 CC improved thrombolytic properties.  
 CC (Updated on 25-MAR-2003 to correct PN field.)  
 XX  
 XX Sequence 2162 BP; 520 A; 606 C; 619 G; 417 T; 0 other;  
 SQ  
 Query Match 94.4%; Score 1065; DB 16; Length 2162;  
 Best Local Similarity 99.5%; Pred. No. 4, 1e-225;  
 Matches 1068; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
 QY 56 CCCAGCGGCTCTGAGGAAACAGTACTGCTACTTTGGGAATGGGTGAGCCTACCGTG 115  
 DB |||||  
 DB 695 CCCTGCTGCTCTGAGGAAACAGTACTGCTACTTTGGGAATGGGTGAGCCTACCGTG 754  
 QY 116 GCACGACAGCTTACCGAGTCGGGTGCTCTGCTCGCTCGGTAATTCATGATCTGA 175  
 DB |||||  
 DB 755 GCACGACAGCTTACCGAGTCGGGTGCTCTGCTCGCTCGGTAATTCATGATCTGA 814

QY 176 TAGCAAGGTTTACACAGCACAGAAACCCAGTGCACAGCAGTGGCCTGGGCAACATA 235  
 DB |||||  
 DB 815 TAGCAAGGTTTACACAGCACAGAAACCCAGTGCACAGCAGTGGCCTGGGCAACATA 874  
 QY 236 ATTACTGCGGAATCTGTATGCGGATGCCAAGCCCTGGTGCACAGTGTGAAGAACCGCA 295  
 DB |||||  
 DB 875 ATTACTGCGGAATCTGTATGCGGATGCCAAGCCCTGGTGCACAGTGTGAAGAACCGCA 934  
 QY 296 GGCTGACGTGGAGTACTGTGATGTGCTCTCCACCTGCGGCTGAGACAGTACA 355  
 DB |||||  
 DB 935 GGCTGACGTGGAGTACTGTGATGTGCTCTCCACCTGCGGCTGAGACAGTACA 994  
 QY 356 GCCAGCCTCAGTTTCGCATCAAAGAGGGCTCTTTGCGGACATGCCTCCACCCCTGGC 415  
 DB |||||  
 DB 995 GCCAGCCTCAGTTTCGCATCAAAGAGGGCTCTTTGCGGACATGCCTCCACCCCTGGC 1054  
 QY 416 AGGTGCTCATCTTTGCCAAGCACAGGAGGTGCGCGGAGAGCGGTCTCTGTGCGGGGCA 475  
 DB |||||  
 DB 1055 AGGTGCTCATCTTTGCCAAGCACAGGAGGTGCGCGGAGAGCGGTCTCTGTGCGGGGCA 1114  
 QY 476 TACTCATCAGCTCTGTGATTTCTCTGCGGCCACTGCTTCCAGGAGGTTTCGCG 535  
 DB |||||  
 DB 1115 TACTCATCAGCTCTGTGATTTCTCTGCGGCCACTGCTTCCAGGAGGTTTCGCG 1174  
 QY 536 CCCACCACTGACGGTGATCTTTGGGCGAAACATACCGGGTGGTCCCTGCGGAGGAGGAGC 595  
 DB |||||  
 DB 1175 CCCACCACTGACGGTGATCTTTGGGCGAAACATACCGGGTGGTCCCTGCGGAGGAGGAGC 1234  
 QY 596 AGAATTTGAAGTCGAAATAATACATTTGTCATAGGAATTCGATGATGACACTTACGACA 655  
 DB |||||  
 DB 1235 AGAATTTGAAGTCGAAATAATACATTTGTCATAGGAATTCGATGATGACACTTACGACA 1294  
 QY 656 ATGACATTTGGCTGCTGCAGCTGAATCGGAATTCGATGATGATGATGATGATGATGATG 715  
 DB |||||  
 DB 1295 ATGACATTTGGCTGCTGCAGCTGAATTCGGAATTCGATGATGATGATGATGATGATG 775  
 QY 716 TGGTCCGCACTGTGTGCTTCCCGCGGAGACCTGCGAGCTGCGGAGCTGGAAGGAGTGTG 775  
 DB |||||  
 DB 1355 TGGTCCGCACTGTGTGCTTCCCGCGGAGACCTGCGAGCTGCGGAGCTGGAAGGAGTGTG 1414  
 QY 776 AGCTCTCGGCTACGGCAAGCATGAGGCTTGTCTCTCTTTCTATTCGAGCGGCTGAAGG 835  
 DB |||||  
 DB 1415 AGCTCTCGGCTACGGCAAGCATGAGGCTTGTCTCTCTTTCTATTCGAGCGGCTGAAGG 1474  
 QY 836 AGGCTCATGTGACAGTGTACCCATCCAGCGCTGCACATCACACATTTACTTAAACAGAA 895  
 DB |||||  
 DB 1475 AGGCTCATGTGACAGTGTACCCATCCAGCGCTGCACATCACACATTTACTTAAACAGAA 1534  
 QY 896 CAGTCACCGCAACATGTGTGCTGAGACACTCGGAGCGGCGGCGCCCGAGGCAAACT 955  
 DB |||||  
 DB 1535 CAGTCACCGCAACATGTGTGCTGAGACACTCGGAGCGGCGGCGCCCGAGGCAAACT 1594  
 QY 956 TGCACGACGCTGCGAGGCGAATTCGGAGGCGCCCTGTGTGTGTGAACATGCGCGCA 1015  
 DB |||||  
 DB 1595 TGCACGACGCTGCGAGGCGAATTCGGAGGCGCCCTGTGTGTGTGAACATGCGCGCA 1654  
 QY 1016 TGACTTTGTGGGCTCATCAGCTGCGGCTGCGGCTGAGACAGAGGATGTCGCGGGTG 1075  
 DB |||||  
 DB 1655 TGACTTTGTGGGCTCATCAGCTGCGGCTGCGGCTGAGACAGAGGATGTCGCGGGTG 1714  
 QY 1076 TGATACAAAGGTTTACCAACTTACCTAGACTGGATTCTGACAAACATGCGACCG 1128  
 DB |||||  
 DB 1715 TGATACAAAGGTTTACCAACTTACCTAGACTGGATTCTGACAAACATGCGACCG 1767

RESULT 10

AAV37294

ID AAV37294 standard; DNA; 2162 BP.

XX AAV37294;

AC AAV37294;

XX 10-SEP-1998 (first entry)

DT 10-SEP-1998 (first entry)

XX DE Human tissue plasminogen activator gene sequence.

XX KW Plasmid pINVI; reverse-splicing intron; group II intron;

XX KW exon binding site; domain V motif; branch site acceptor;

XX KW nucleophilic group; transesterification; phosphodiester bond;

XX KW autocatalytic Y-branched intron; reverse splicing reaction; ds.

XX OS Homo sapiens.

XX FH Key

XX FT misc\_feature 82..334 Location/Qualifiers

XX FT /tag= a

XX FT /note= "signal sequence and finger-like domain"

XX FT misc\_feature 335..447

XX FT /tag= b

XX FT /note= "EGF-like domain"

XX FT misc\_feature 448..714

XX FT /tag= c

XX FT /note= "Kringle-1 domain"

XX FT misc\_feature 715..972

XX FT /tag= d

XX FT /note= "Kringle-2 domain"

XX FT misc\_feature 973..2162

XX FT /tag= e

XX FT /note= "catalytic domain"

XX PN US5780272-A.

XX XX 14-JUL-1998.

XX XX 07-JUN-1995; 95US-0488015.

XX XX 10-SEP-1993; 93US-0119512.

XX XX (HARD ) HARVARD COLLEGE.

XX XX Jarrell KA;

XX XX WPI; 1998-413060/35.

XX XX Reverse splicing construct containing fragments of autocatalytic

XX XX introns - able to cleave and ligate discontinuous nucleic acid for

XX XX generating new genes and e.g. ribozymes, libraries of enzymes and

XX XX antibodies

XX PS Example 4; Columns 53-56; 56pp; English.

XX CC The present sequence represents the human tissue plasminogen activator

XX CC gene. It was used to construct plasmid TPA-KS+, which is used in the

XX CC course of the invention. The specification describes a purified

XX CC reverse-splicing intron which comprises a segment comprising a

XX CC 5'-part of a group II intron, including an exon binding site not

XX CC naturally present in the intron and a second segment comprising a

XX CC 3'-part of a group II intron, including a domain V motif, a branch

XX CC site acceptor, and a nucleophilic group for transesterifying a

XX CC phosphodiester bond of an RNA. Together the two segments form an

XX CC autocatalytic Y-branched intron which catalyses integration of at least

XX CC the first segment into substrate RNA by a reverse splicing reaction

XX CC The reverse-splicing introns are used, by specific cleavage and ligation

XX CC of discontinuous nucleic acid, to generate new genes and gene products,

XX CC e.g. ribozymes (for use in gene therapy or as reagents in DNA

XX CC manipulation, e.g. replacements for restriction enzymes) or

XX CC immunologically active or signal-transducing proteins such as antibody

XX CC and enzyme libraries.

XX SQ Sequence 2162 BP; 520 A; 606 C; 619 G; 417 T; 0 other;

XX Query Match 94.4%; Score 1065; DB 19; Length 2162;

XX Best Local Similarity 99.5%; Pred. No. 4.1e-22;

XX Matches 1068; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

XX 56 CCCAGCGCGCTCTGAGGGAACAGTGACTGTCTTGTGGGAATGGGTACGCTACCGTG 115

Db 695 CCCCTGCTGCTCTGAGGGAACAGTGACTGTCTTGTGGGAATGGGTACCGTG 754

Qy 116 GCAGCAGAGCTCAGCGAGTGGGTGCTCTGCTCCCGTGGAAATCCATGATCTCTGA 175

Db 755 GCAGCAGAGCTCAGCGAGTGGGTGCTCTGCTCCCGTGGAAATCCATGATCTCTGA 814

Qy 176 TAGGCAAGGTTTACACAGCAGAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 235

Db 815 TAGGCAAGGTTTACACAGCAGAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 874

Qy 236 ATTAAGTCCGGAATCTCTGATGGGATGCCAAGCCCTGGTGCACAGTGTCTGAAGAACCGCA 295

Db 875 ATTAAGTCCGGAATCTCTGATGGGATGCCAAGCCCTGGTGCACAGTGTCTGAAGAACCGCA 934

Qy 296 GGCTGAGCTGGGAGTACTGTGATGTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTG 355

Db 935 GGCTGAGCTGGGAGTACTGTGATGTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTG 994

Qy 356 GCCAGCTCTAGTTCGATCAAAAGGAGGGTCTTTTCGCCGACATCGCTCCCGCCCTGGC 415

Db 995 GCCAGCTCTAGTTCGATCAAAAGGAGGGTCTTTTCGCCGACATCGCTCCCGCCCTGGC 1054

Qy 416 AGGCTGCCATCTTTGCCAAGCAGAGAGGTGCGCCGAGAGCGGTTCCTGTGCGGGGCA 475

Db 1055 AGGCTGCCATCTTTGCCAAGCAGAGAGGTGCGCCGAGAGCGGTTCCTGTGCGGGGCA 1114

Qy 476 TACTCATGCTCTGCTGATCTCTGCTGCGCCCACTGCTTCAGAGAGGTTTCGCG 535

Db 1115 TACTCATGCTCTGCTGATTTCTCTGCGGCCACTGCTTCAGAGAGGTTTCGCG 1174

Qy 536 CCCACCACTGAGCGTGATCTTGGGAGAAACATACCGGTGTCCTCCGCGAGAGGAGC 595

Db 1175 CCCACCACTGAGCGTGATCTTGGGAGAAACATACCGGTGTCCTCCGCGAGAGGAGC 1234

Qy 596 AGAAATTTGAAGTCGAAAAATACATTTGTCCATAGGAATTCGATGATGACACTTTACGACA 655

Db 1235 AGAAATTTGAAGTCGAAAAATACATTTGTCCATAGGAATTCGATGATGACACTTTACGACA 1294

Qy 656 ATGACATTTGGCTGCTGAGCTGAAATCGGATTCGTCCTGCTGTCGCCAGAGAGCAGCG 715

Db 1295 ATGACATTTGGCTGCTGAGCTGAAATCGGATTCGTCCTGCTGTCGCCAGAGAGCAGCG 1354

Qy 716 TGGTCCGACATGTGTGCTTCCCGCGGAGACCTGAGCTGCGGAGCTGGAACGAGTGTG 775

Db 1355 TGGTCCGACATGTGTGCTTCCCGCGGAGACCTGAGCTGCGGAGCTGGAACGAGTGTG 1414

Qy 776 AGCTCTCCGGCTACGGCAAGCATGAGGCTTGTCTCTCTTTCTATTTCGAGCGGCTGAAGG 835

Db 1415 AGCTCTCCGGCTACGGCAAGCATGAGGCTTGTCTCTCTTTCTATTTCGAGCGGCTGAAGG 1474

Qy 836 AGGCTCATGTGAGTGTATCCATTCAGCGCTGCTGCACATCAACATTTACTTTACAGAA 895

Db 1475 AGGCTCATGTGAGTGTATCCATTCAGCGCTGCTGCACATCAACATTTACTTTACAGAA 1534

Qy 896 CAGTCACGCAACATGCTGTGCTGAGACACTCGAGCGCGGCGGCCAGAGCAACT 955

Db 1535 CAGTCACGCAACATGCTGTGCTGAGACACTCGAGCGCGGCGGCCAGAGCAACT 1594

Qy 956 TGCAACGAGCGCTGCGAGGCGATTCGGGAGCGCCCTGCTGTGTGTCTGAACGATGCGCGCA 1015

Db 1595 TGCAACGAGCGCTGCGAGGCGATTCGGGAGCGCCCTGCTGTGTGTCTGAACGATGCGCGCA 1654

Qy 1016 TGACTTTGGTGGGATCATGAGTGGGCGCTGCGGCTGTGAGCAGAGATGTCGCGGTG 1075

Db 1655 TGACTTTGGTGGGATCATGAGTGGGCGCTGCGGCTGTGAGCAGAGATGTCGCGGTG 1714

Qy 1076 TGATACAAAGGTTTACCAACTTACCTAGACTGGATTCTGTGACCAATCGACCG 1128

Db 1715 TGATACAAAGGTTTACCAACTTACCTAGACTGGATTCTGTGACCAATCGACCG 1767

RESULT 11





OS Homo sapiens.  
 XX Key Location/Qualifiers  
 FH sig\_peptide 102..206  
 FT /\*tag= a  
 FT mat\_peptide ..1787  
 FT /\*tag= b  
 FT /product=t-PA  
 XX EP297066-A.  
 XX 28-DEC-1988.  
 PD 14-JUN-1988; 88EP-0850207.  
 XX 18-JUN-1987; 87SE-0002562.  
 XX (KABI ) KABIGEN AB.  
 XX Pohl G, Hansson L, Loewenadler B;  
 XX WPI: 1989-001503/01.  
 DR P-PSDB; AAF93716.  
 XX Modified tissue plasminogen-activator - having domains deleted and amino acid changes to increase biological half-life and reduce inactivation.  
 XX Disclosure; Fig 1; 18pp; English.  
 CC The sequence was obtained from cDNA prep. from mRNA extracted from Bowes melanoma cells. The cDNA was used to construct a gene library which was screened with a partial t-PA cDNA clone. Plasmid pKG12 was isolated and shown to comprise the whole coding region for human t-PA as well as 102bp 5'flanking, 760 bp 3' flanking DNA and a poly A tail.  
 CC See also AAN91608.  
 CC (Updated on 25-MAR-2003 to correct PI field.)  
 XX Sequence 2560 BP; 634 A; 690 C; 689 G; 547 T; 0 other;  
 SQ Query Match 94.4%; Score 1065; DB 10; Length 2560;  
 Best Local Similarity 99.5%; Pred. No. 4.2e-225;  
 Matches 1068; Conservative 0; Mismatches 25; Indels 0; Gaps 0;  
 OY 56 CCCAGCGCCCTCTGAGGAAACAGTACTGTCTACTTTGGGAATGGGTACGCTACCGTG 115  
 DB 715 CCCCTGCTCTCTGAGGAAACAGTACTGTCTACTTTGGGAATGGGTACGCTACCGTG 774  
 OY 116 GCACGACAGCCTCAACGAGTCGGTGCTCTGCTCCCGTGGAAATCCATGATPCTGA 175  
 DB 775 GCACGACAGCCTCAACGAGTCGGTGCTCTGCTCCCGTGGAAATCCATGATPCTGA 834  
 OY 176 TAGGCAAGGTTTACACACACAGACCCAGTCCAGGCACTGGGCTGGGCAACATA 235  
 DB 835 TAGGCAAGGTTTACACACACAGACCCAGTCCAGGCACTGGGCTGGGCAACATA 894  
 OY 236 ATTACTGCGGGAATCTGATGGGATGCCAAGCCCTGGTGCCAGTCTGAAGAACCGCA 295  
 DB 895 ATTACTGCGGGAATCTGATGGGATGCCAAGCCCTGGTGCCAGTCTGAAGAACCGCA 954  
 OY 296 GGCTGACGTGGGAGTACTGTGATGTGCCCTCTGCTCCACCTGGGCTGGAGACAGTACA 355  
 DB 955 GGCTGACGTGGGAGTACTGTGATGTGCCCTCTGCTCCACCTGGGCTGGAGACAGTACA 1014  
 OY 356 GCCAGCCTCAGTTTGGCATCAAGAGAGGCTCTTGGCCGACATGCCCTCCACCCCTGGC 415  
 DB 1015 GCCAGCCTCAGTTTGGCATCAAGAGAGGCTCTTGGCCGACATGCCCTCCACCCCTGGC 1074  
 OY 416 AGGCTGCCATCTTTGCCAAGACACAGAGGTGCGCCGGAGAGCGGTCTCTGTGCGGGGCA 475  
 DB 1075 AGGCTGCCATCTTTGCCAAGACACAGAGGTGCGCCGGAGAGCGGTCTCTGTGCGGGGCA 1134  
 OY 476 TACTCATCAGTCTCTGTGGATTCTCTCTGCGGCCCACTGTCTTCAGAGAGGTTTCGCG 535

DB 1135 TACTCATCAGTCTCTGTGGATTCTCTCTGCGGCCACTGTCTCCAGAGAGGTTTCGCG 1194  
 OY 536 CCACACACCTGAGCGTGATCTTTGGGAGAAACATACCGGCTGGTCCCTGGGAGAGAGC 595  
 DB 1195 CCACACACCTGAGCGTGATCTTTGGGAGAAACATACCGGCTGGTCCCTGGGAGAGAGC 1254  
 OY 596 AGAAATTTGAAGTCGAAAAATACATTGTCTCCATAGGAATTCGATGATGACACTTTACGACA 655  
 DB 1255 AGAAATTTGAAGTCGAAAAATACATTGTCTCCATAGGAATTCGATGATGACACTTTACGACA 1314  
 OY 656 ATGACATTTGGCTGCTGCTGAGCTGAATTCGATTCGCTCCGCTGTGCCAGAGAGCAGCG 715  
 DB 1315 ATGACATTTGGCTGCTGCTGAGCTGAATTCGATTCGCTCCGCTGTGCCAGAGAGCAGCG 1374  
 OY 716 TGGTCCGCACTGTGTGCTTTCCCGCGGACCTGCGAGTCCGAGTCTGACCGAGTGTG 775  
 DB 1375 TGGTCCGCACTGTGTGCTTTCCCGCGGACCTGCGAGTCCGAGTCTGACCGAGTGTG 1434  
 OY 776 AGCTCTCCGCTACGGCAAGCATGAGGCTTTGTCTCTCTTCTATTTCGAGCGGCTGAAGG 835  
 DB 1435 AGCTCTCCGCTACGGCAAGCATGAGGCTTTGTCTCTTCTATTTCGAGCGGCTGAAGG 1494  
 OY 836 AGGCTCATGTGCACTGTACCCATCCATCCAGCGCTGCAATCACAAATTTACTTAACAGAA 895  
 DB 1495 AGGCTCATGTGCACTGTACCCATCCATCCAGCGCTGCAATCACAAATTTACTTAACAGAA 1554  
 OY 896 CAGTCACCGACACATGCTGTGCTGAGACACATCGGAGCGGCGGCCAGGCAACT 955  
 DB 1555 CAGTCACCGACACATGCTGTGCTGAGACACATCGGAGCGGCGGCCAGGCAACT 1614  
 OY 956 TGCAACGAGCGCTGCGAGGCGGATTCGAGGAGCGGCCCTTGCTGTCTGAAACGATGCGCGCA 1015  
 DB 1615 TGCAACGAGCGCTGCGAGGCGGATTCGAGGAGCGGCCCTTGCTGTCTGAAACGATGCGCGCA 1674  
 OY 1016 TGACTTTGGTGGGCAATCATCAGCTGGGCGCTGCGGCTGTGACAGAGATGTCGCGGCTG 1075  
 DB 1675 TGACTTTGGTGGGCAATCATCAGCTGGGCGCTGCGGCTGTGACAGAGATGTCGCGGCTG 1734  
 OY 1076 TGACACAAAGGTTTACCAACTTACCTAGACTGGATTGCTGACAAACATCGGACCG 1128  
 DB 1735 TGACACAAAGGTTTACCAACTTACCTAGACTGGATTGCTGACAAACATCGGACCG 1787  
 RESULT 13  
 AAN70491  
 ID AAN70491 standard; cDNA; 7533 BP.  
 XX AAN70491;  
 AC 25-MAR-2003 (updated)  
 XX 31-OCT-2002 (updated)  
 DT 07-JUN-1991 (first entry)  
 DT  
 XX  
 DE Entire sequence of high level expression vector pEm1-tpa for tissue plasminogen activator (TPA).  
 DE Enhancer-blocking element; expression vector; ss.  
 KW Unidentified.  
 XX  
 OS  
 XX Key Location/Qualifiers  
 FH mRNA 20..1708  
 FT /\*tag= a  
 FT misc\_feature 7..1748  
 FT /\*tag= b  
 FT /label= TPA cDNA  
 FT 1749..1955  
 FT 3'UTR /\*tag= c  
 FT /note= "from immunoglobulin kappa light chain gene"  
 FT 1956..2164  
 FT 3'clip /\*tag= d  
 FT /note= "as above"  
 FT misc\_feature 2165..4463

FT /tag= e  
 FT /label= pBR322  
 FT /note= "contg. the ampicillinase gene and the  
 FT bacterial origin of replication"  
 FT misc\_feature complement (4464..6436)  
 FT /tag= f  
 FT /label= E. coli gpt gene  
 FT 6437..6547  
 FT /tag= g  
 FT /label= SV40 origin region  
 FT 6454..6468  
 FT /tag= h  
 FT /label= SV40 (gpt) CAP site  
 FT 6548..6715  
 FT /tag= i  
 FT /label= SV40 enhancer (TZ bp repeat)  
 FT 6716..6971  
 FT /tag= j  
 FT /label= lambda-1  
 FT /note= "functions as a blocking element"  
 FT 6972..7281  
 FT /tag= k  
 FT /note= "from immunoglobulin heavy chain (IGH)"  
 FT 7282..7466  
 FT /tag= l  
 FT /note= "from metallothioneine 1 (MT)"  
 FT 7467..7533  
 FT /tag= m  
 FT /label= MT CAP site  
 FT 1930..1935  
 FT /tag= n  
 FT polyA\_signal  
 FT  
 XX  
 PN EP237157-A.  
 XX  
 PD 16-SEP-1987.  
 XX  
 XX 27-JAN-1987; 87EP-0300658.  
 PF  
 XX 07-MAR-1986; 86US-0837595.  
 PR  
 XX (DAMO-) DAMON BIOTECH INC.  
 PA (ABBO) ABBOTT BIOTECH INC.  
 PA  
 PI Gillies SD;  
 XX  
 DR WPI; 1987-258421/37.  
 XX  
 XX DNA construct for high level expression of protein - having a  
 PT blocking element so that marker protein is produced only at  
 PT levels required for selection  
 XX  
 PS Disclosure; Fig. 9A-D; 20pp; English.  
 XX  
 CC Vector pEM1-tpa is an example of a vector of the invention. Pref.  
 CC the blocking element is a promoter sequence oriented with its native  
 CC 3' end disposed proximal to the enhancer element and its 5' end  
 CC disposed distal to the enhancer element. Alternatively the  
 CC interposed blocking element comprises a promoter sequence with its 5'  
 CC end disposed proximal to the enhancer element and, at its 3' end, a  
 CC gene encoding for a different protein of interest, (see FT). The  
 CC vector pEM1 was constructed by inserting an enhancer-blocking  
 CC element into pEM1.  
 CC TPA cDNA fragment was inserted into the unique XhoI site present in  
 CC both pEM1 and pEM1. Of 26 transformants obtd. with pEM1-tpa, 5  
 CC produced TPA at levels of 140-500 IU/ml. Of 16 transformants obtd.  
 CC with pEM1-tpa, 7 produced TPA at levels of 1000-6000 IU/ml.  
 CC (Updated on 31-OCT-2002 to add missing OS field.)  
 CC (Updated on 25-MAR-2003 to correct PA field.)  
 XX  
 SQ Sequence 7533 BP; 1921 A; 1899 C; 1844 G; 1868 T; 1 other;  
 Query Match 94.4%; Score 1065; DB 8; Length 7533;  
 Best Local Similarity 99.5%; Pred. No. 4.8e-225;

Matches 1068; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
 Qy 56 CCCAGGGCGCCTCTGAGGGAACAGTACTACTTGTGGAAATGGGTTCAGCTACCGTG 115  
 Db |||||  
 Qy 633 CCCCTGCTGCTCTGAGGGAACAGTACTACTTGTGGAAATGGGTTCAGCTACCGTG 692  
 Db |||||  
 Qy 116 GCACGCACAGCTCACCGAGTCCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTGA 175  
 Db |||||  
 Qy 693 GCACGCACAGCTCACCGAGTCCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTGA 752  
 Db |||||  
 Qy 176 TAGGCAAGGTTTACACAGCAGAACCCAGTCCAGGACACTGGGCTGGGCAACATA 235  
 Db |||||  
 Qy 753 TAGGCAAGGTTTACACAGCAGAACCCAGTCCAGGACACTGGGCTGGGCAACATA 812  
 Db |||||  
 Qy 236 ATTACTCCGGAATCTCTGATGGGATGCCAAGCCCTGGTGCACGTCTGAAGAACCGCA 295  
 Db |||||  
 Qy 813 ATTACTCCGGAATCTCTGATGGGATGCCAAGCCCTGGTGCACGTCTGAAGAACCGCA 872  
 Db |||||  
 Qy 296 GGCTGAGTGGGAGTACTGTGATGTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTGACAGTACA 355  
 Db |||||  
 Qy 873 GGCTGAGTGGGAGTACTGTGATGTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTGACAGTACA 932  
 Db |||||  
 Qy 356 GCCAGCCTCAGTTTCGATCAAGAGGAGGCTCTTTCGCGACATCGCTCCACCCCTGGC 415  
 Db |||||  
 Qy 933 GCCAGCCTCAGTTTCGATCAAGAGGAGGCTCTTTCGCGACATCGCTCCACCCCTGGC 992  
 Db |||||  
 Qy 416 AGGCTGCCATCTTTGCCAAGCACAGGAGGTCCGCCGAGAGCGGTTCCTGTGCGGGGCA 475  
 Db |||||  
 Qy 993 AGGCTGCCATCTTTGCCAAGCACAGGAGGTCCGCCGAGAGCGGTTCCTGTGCGGGGCA 1052  
 Db |||||  
 Qy 476 TACTCATCAGCTCTCTGATTTCTCTGCGGCCACTGCTTCCAGAGAGGTTTCGC 535  
 Db |||||  
 Qy 1053 TACTCATCAGCTCTCTGATTTCTCTGCGGCCACTGCTTCCAGAGAGGTTTCGC 1112  
 Db |||||  
 Qy 536 CCCACCCTGACCGTCTTGGGAGACATACCGGTGGTCCCTGGCGAGGAGGAGC 595  
 Db |||||  
 Qy 1113 CCCACCCTGACCGTCTTGGGAGACATACCGGTGGTCCCTGGCGAGGAGGAGC 1172  
 Db |||||  
 Qy 596 AGAAATTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA 655  
 Db |||||  
 Qy 1173 AGAAATTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACA 1232  
 Db |||||  
 Qy 656 ATGACATTGGCGCTCTGACGCTGAAATCGGATTCGTCCTGTCGCCAGAGAGCAGCG 715  
 Db |||||  
 Qy 1233 ATGACATTGGCGCTCTGACGCTGAAATCGGATTCGTCCTGTCGCCAGAGAGCAGCG 1292  
 Db |||||  
 Qy 716 TGGTCCGACATGTGCTCTTCCCGGGGACCTGCGAGCTGCCGACCTGGAGCGAGTGTG 775  
 Db |||||  
 Qy 1293 TGGTCCGACATGTGCTCTTCCCGGGGACCTGCGAGCTGCCGACCTGGAGCGAGTGTG 1352  
 Db |||||  
 Qy 776 AGCTCTCCGCTACGGCAAGCATGAGGCTTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTGAAGG 835  
 Db |||||  
 Qy 1353 AGCTCTCCGCTACGGCAAGCATGAGGCTTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTGAAGG 1412  
 Db |||||  
 Qy 836 AGGCTCATGTGACAGTATCCCATCCAGCGCTGCACATCAACAATTTACTTAACAGAA 895  
 Db |||||  
 Qy 1413 AGGCTCATGTGACAGTATCCCATCCAGCGCTGCACATCAACAATTTACTTAACAGAA 1472  
 Db |||||  
 Qy 896 CAGTCACGCAACAATCTGTGTGCTGAGACACTCGAGCGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 955  
 Db |||||  
 Qy 1473 CAGTCACGCAACAATCTGTGTGCTGAGACACTCGAGCGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 1532  
 Db |||||  
 Qy 956 TGACAGACGCTGCGAGCGGATTCGGAGAGCGCCCTGCTGTGTGTCTGAACATGATGGCGCA 1015  
 Db |||||  
 Qy 1533 TGACAGACGCTGCGAGCGGATTCGGAGAGCGCCCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1592  
 Db |||||  
 Qy 1016 TGACTTTGGTGGGCATCATCAGCTGGGGCCCTGGGCTGTGGACAGAGGATGTCCCGGGTG 1075  
 Db |||||  
 Qy 1593 TGACTTTGGTGGGCATCATCAGCTGGGGCCCTGGGCTGTGGACAGAGGATGTCCCGGGTG 1652  
 Db |||||  
 Qy 1076 TGACACAAAGGTTTACCAACTACCTAGACTGGATTCTGTGACAAATGGGACCG 1128  
 Db |||||  
 Qy 1653 TGACACAAAGGTTTACCAACTACCTAGACTGGATTCTGTGACAAATGGGACCG 1705  
 Db |||||



QY 416 AGGCTGCCATCTTTGCAAGCAAGAGGTGCCCCGGAAGCGGTTCTGTGCGGGGCA 475  
 Db |||||  
 QY 993 AGGCTGCCATCTTTGCAAGCAAGAGGTGCCCCGGAAGCGGTTCTGTGCGGGGCA 1052  
 Db |||||  
 QY 476 TACTCATCAGCTCTGCTGGATTCTCTGCGCGCCACTGCTTCCAGGAGGTTTCCGC 535  
 Db |||||  
 QY 1053 TACTCATCAGCTCTGCTGGATTCTCTGCGCGCCACTGCTTCCAGGAGGTTTCCGC 1112  
 Db |||||  
 QY 536 CCCACCACTGACGCTGATCTTTGGCAGAACATACCGGCTGCTTCCAGGAGGAGC 595  
 Db |||||  
 QY 1113 CCCACCACTGACGCTGATCTTTGGCAGAACATACCGGCTGCTTCCAGGAGGAGC 1172  
 Db |||||  
 QY 596 AGAAATTTGAAGTCGAAATATCATTTGCCATAAGAAATTCGATGATGACACTTACGACA 655  
 Db |||||  
 QY 1173 AGAAATTTGAAGTCGAAATATCATTTGCCATAAGAAATTCGATGATGACACTTACGACA 1232  
 Db |||||  
 QY 656 ATGACATTTGCGCTGCTGACGCTGAATCGGATTGCTGCGGCTGCTGCCAGGAGACGG 715  
 Db |||||  
 QY 1233 ATGACATTTGCGCTGCTGACGCTGAATCGGATTGCTGCGGCTGCTGCCAGGAGACGG 1292  
 Db |||||  
 QY 716 TGGTCCGCACTGTGTGCTTCCCGCGGACCTGCACTGCGGACCTGGACGGAGTGTG 775  
 Db |||||  
 QY 1293 TGGTCCGCACTGTGTGCTTCCCGCGGACCTGCACTGCGGACCTGGACGGAGTGTG 1352  
 Db |||||  
 QY 776 AGCTCTCGGCTACCGGAGCATGAGGCTTGTCTCTTCTTATTTGCGAGCGGCTGAAGG 835  
 Db |||||  
 QY 1353 AGCTCTCGGCTACCGGAGCATGAGGCTTGTCTCTTCTTATTTGCGAGCGGCTGAAGG 1412  
 Db |||||  
 QY 836 AGGCTCATGTGACGCTGACCATCCAGCCGCTGCACATCACATTTACTTTACAGAA 895  
 Db |||||  
 QY 1413 AGGCTCATGTGACGCTGACCATCCAGCCGCTGCACATCACATTTACTTTACAGAA 1472  
 Db |||||  
 QY 896 CAGTCACCGACAACTGCTGTGTGCTGGAGACACTCGGAGCGCGGCGCCAGGCAACT 955  
 Db |||||  
 QY 1473 CAGTCACCGACAACTGCTGTGTGCTGGAGACACTCGGAGCGCGGCGCCAGGCAACT 1532  
 Db |||||  
 QY 956 TGCAGAGCGCTGCCAGGCGGATTCGGAGGCGCCCTGCTGTGTGTGCTGAACGATGCCGCA 1015  
 Db |||||  
 QY 1533 TGCAGAGCGCTGCCAGGCGGATTCGGAGGCGCCCTGCTGTGTGTGCTGAACGATGCCGCA 1592  
 Db |||||  
 QY 1016 TGACTTTGTTGGGATCATACGCTGCGGCGCTGGGCTGTGGACAGAGGATGTCGCGGTG 1075  
 Db |||||  
 QY 1593 TGACTTTGTTGGGATCATACGCTGCGGCGCTGGGCTGTGGACAGAGGATGTCGCGGTG 1652  
 Db |||||  
 QY 1076 TGTACAAAAGGTTACCACTACCTAGACTGGATTGCTGACAACTGCGACCG 1128  
 Db |||||  
 QY 1653 TGTACAAAAGGTTACCACTACCTAGACTGGATTGCTGACAACTGCGACCG 1705  
 Db |||||

## RESULT 15

ID AAT27588  
 AC AAT27588 standard; DNA; 1680 BP.  
 AC AAT27588;  
 DT 25-MAR-2003 (updated)  
 DT 06-AUG-1996 (first entry)  
 XX Novel plasminogen activator DNA.

XX Tissue plasminogen activator; tPA; alpha2-plasmin inhibitor;  
 KW fibrinolytic; thrombolytic; fibrin; thrombosis; blood clotting;  
 KW protein engineering; kringle; ss.  
 XX Synthetic.

XX Location/Qualifiers  
 FH sig\_peptide 1..105  
 FT /tag= a  
 FT mat\_peptide 106..1677  
 FT /tag= b

PN US5504001-A.  
 XX 02-APR-1996.  
 XX 06-JUN-1994; 94US-0254485.  
 XX 25-NOV-1987; 87US-0125629.  
 PR 28-JAN-1992; 92US-0827587.  
 PR 06-JUN-1994; 94US-0254485.  
 XX (ZYMO) ZYMOGENETICS INC.  
 XX Foster DC;  
 XX WPI; 1996-187699/19.  
 DR P-PSDB; AAR96223.  
 XX Hybrid plasminogen activator comprises human tPA activator and  
 PT N-terminal crosslinking domain from alpha2-plasmin inhibitor -  
 PT useful to treat thrombosis and image blood clots  
 XX Example 3; Fig 11; 35pp; English.  
 CC A DNA construct (AAT27586) codes for a novel plasminogen  
 CC activator (AAR96222) in which the kringle K1 domain (AAR96221)  
 CC of plasminogen replaces the native K1 domain of tissue  
 CC plasminogen activator (tPA) (AAR96220). It was obt. by  
 CC inserting a synthetic plasminogen K1 domain DNA (AAT27586)  
 CC into tPA cDNA and can be expressed in transformed Escherichia  
 CC coli RRI cells (FERM P-9272). Novel plasminogen activators are  
 CC produced that show increased clot lysing specificity or plasma  
 CC half-life.  
 CC (Updated on 25-MAR-2003 to correct PF field.)  
 XX SQ Sequence 1680 BP; 387 A; 471 C; 479 G; 343 T; 0 other;

Query Match 94.3%; Score 1064; DB 17; Length 1680;  
 Best Local Similarity 99.5%; Pred. No. 6.6e-225;  
 Matches 1067; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 57 CCAGCGCGCTCTGAGGGAACAGTGACTGCTCTTTGGGAATGGGTACGCTACCGTGG 116  
 Db |||||  
 QY 606 CTTGGAATGCTCTGAGGGAACAGTGACTGCTCTTTGGGAATGGGTACGCTACCGTGG 665  
 Db |||||  
 QY 117 CACGACAGGCTCACGAGTGGGTGCTCTCTGCTCCCGTGAATTCATGATCTCGAT 176  
 Db |||||  
 QY 666 CACGACAGGCTCACGAGTGGGTGCTCTCTGCTCCCGTGAATTCATGATCTCGAT 725  
 Db |||||  
 QY 177 AGGCAAGGTTTACACAGACAGAACCCGAGTGCAGGCACTGGGCTGGGCAACATAA 236  
 Db |||||  
 QY 726 AGGCAAGGTTTACACAGACAGAACCCGAGTGCAGGCACTGGGCTGGGCAACATAA 785  
 Db |||||  
 QY 237 TTTACTGCGGAATCTGATGGGATGCAAGCCCTGGTGCACGTGTGAGRACCCGAG 296  
 Db |||||  
 QY 786 TTTACTGCGGAATCTGATGGGATGCAAGCCCTGGTGCACGTGTGAGRACCCGAG 845  
 Db |||||  
 QY 297 GCTGACGTGGGAGTACTGTGATGTGCCCTCTCTGCTCCACCTGCGGCTTGAGACAGTACAG 356  
 Db |||||  
 QY 846 GCTGACGTGGGAGTACTGTGATGTGCCCTCTCTGCTCCACCTGCGGCTTGAGACAGTACAG 905  
 Db |||||  
 QY 357 CCAGCCTCAGTTTGGCATCAAGAGGCTCTTGGCCGACATCGCTCCCAACCCCTGGCA 416  
 Db |||||  
 QY 906 CCAGCCTCAGTTTGGCATCAAGAGGCTCTTGGCCGACATCGCTCCCAACCCCTGGCA 965  
 Db |||||  
 QY 417 GGCTGCCATCTTTGCCAAGCAGAGGAGTGCAGGAGGCTTCTGTGCGGGGCGAT 476  
 Db |||||  
 QY 966 GGCTGCCATCTTTGCCAAGCAGAGGAGTGCAGGAGGCTTCTGTGCGGGGCGAT 1025  
 Db |||||  
 QY 477 ACTCATCAGTCTCTGTGGATTCTCTGCGCGCCACTGCTTCCAGAGAGGTTTCCGCC 536  
 Db |||||  
 QY 1026 ACTCATCAGTCTCTGTGGATTCTCTGCGCGCCACTGCTTCCAGAGAGGTTTCCGCC 1085  
 Db |||||  
 QY 537 CCACCACCTGACGCTGATCTTTGGGCAACATACCGGGTGTCTCTGCGGAGGAGGAGCA 596  
 Db |||||



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 18, 2003, 22:51:23 ; Search time 88.2804 Seconds  
(without alignments)  
5639.757 Million cell updates/sec

Title: US-09-987-455-2

Perfect score: 1128

Sequence: 1 atgaaaagacagctatcgc.....ttctgacacatcgacg 1128

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA.\*

- 1: /cgn2\_6/ptodata/2/ina/5A\_COMB.seq.\*
- 2: /cgn2\_6/ptodata/2/ina/5B\_COMB.seq.\*
- 3: /cgn2\_6/ptodata/2/ina/6A\_COMB.seq.\*
- 4: /cgn2\_6/ptodata/2/ina/6B\_COMB.seq.\*
- 5: /cgn2\_6/ptodata/2/ina/PCTUS\_COMB.seq.\*
- 6: /cgn2\_6/ptodata/2/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1065	94.4	1419	2	US-08-811-949-62
2	1065	94.4	1848	3	US-08-814-412-10
3	1065	94.4	2101	2	US-08-811-949-42
4	1065	94.4	2162	1	US-08-119-512-3
5	1065	94.4	2162	1	US-08-488-015B-3
6	1065	94.4	2162	1	US-08-488-015B-25
7	1063.4	94.3	1314	2	US-08-811-949-48
8	1063.4	94.3	1738	6	5200340-1
9	1063.4	94.3	1955	2	US-08-883-795A-39
10	1063.4	94.3	2457	6	5344773-1
11	1063.4	94.3	7360	1	US-08-286-740-1
12	1063.4	94.3	7360	5	PCT-US95-09576-1
13	1062	94.1	1068	2	US-08-811-949-44
14	1061.8	94.1	1314	2	US-08-811-949-50
15	1061.8	94.1	1974	2	US-08-811-949-38
16	1060.4	94.0	1170	2	US-08-811-949-64
17	1058.8	93.9	1068	2	US-08-811-949-46
18	1057.9	93.7	1314	2	US-08-811-949-54
19	1056.4	93.7	1068	1	US-08-137-116-2
20	1054.8	93.5	1065	1	US-08-427-640-1
21	1054.8	93.5	1137	4	US-09-553-498-9
22	1054.8	93.5	1137	4	US-09-618-869-9
23	1054.8	93.4	1068	2	US-08-811-949-52
24	1053.8	93.4	1314	2	US-08-811-949-56
25	1052.4	93.3	1170	2	US-08-811-949-66
26	1050	93.1	1068	6	5223256-3
27	1046.8	92.8	1065	1	US-08-427-640-5

28	1038	92.0	1068	2	US-08-811-949-58
29	1035.4	91.8	1068	1	US-08-427-640-3
30	1030	91.3	1065	2	US-08-811-949-60
31	953.4	84.5	2497	6	5185259-2
32	924	81.9	1068	1	US-08-427-640-7
33	798	70.7	1163	2	US-08-558-269-5
34	798	70.7	1163	3	US-09-410-882-5
35	625.8	55.5	1727	6	5244676-4
36	625.6	55.5	1605	6	5244676-1
37	472	41.8	472	2	US-08-811-949-40
38	414	36.7	453	6	5200340-3
39	267.4	23.7	1724	6	5200340-5
40	170.8	15.1	1236	1	US-07-957-039A-7
41	170.8	15.1	1475	4	US-09-643-597-122
42	170.8	15.1	1475	4	US-09-480-884A-122
43	170.8	15.1	1475	4	US-09-542-615A-122
44	170.8	15.1	1475	4	US-09-606-421B-122
45	170.8	15.1	2301	6	5188829-2

ALIGNMENTS

RESULT 1

US-08-811-949-62  
; Sequence 62, Application US/08811949  
; Patent No. 5840533  
; GENERAL INFORMATION:  
; APPLICANT: NIWA, MINEO  
; APPLICANT: SAITO, YOSHIMASA  
; APPLICANT: SASAKI, HITOSHI  
; APPLICANT: HAYASHI, MASAKO  
; APPLICANT: NOTANI, JOUJI  
; APPLICANT: KOBAYASHI, MASAKAZU  
; TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR  
; NUMBER OF SEQUENCES: 67  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,  
; ADDRESSEE: P.C.  
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400  
; CITY: ARLINGTON  
; STATE: VA  
; COUNTRY: USA  
; ZIP: 22202

COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/811,949  
; FILING DATE: 05-MAR-1997  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: OBLON, NORMAN F.  
; REGISTRATION NUMBER: 24,618  
; REFERENCE/DOCKET NUMBER: 18-966-0  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 703-413-3000  
; TELEFAX: 703-413-2220  
; INFORMATION FOR SEQ ID NO: 62:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1419 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: circular  
; MOLECULE TYPE: DNA (genomic)  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 1..1416

US-08-811-949-62  
Query Match 94.4%; Score 1065; DB 2; Length 1419;

Best Local Similarity 99.5%; Pred. No. 7.7e-290;  
Matches 1068; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 56 CCCAGGCGCCTCTGAGGAAACAGTGTACTTTGGGAATGGGTACGCTTACCGTG 115  
Db |||||  
344 CCCCTGCCCTGCTCTGAGGAAACAGTGTACTTTGGGAATGGGTACGCTTACCGTG 403  
QY 116 GCACGACAGCCTTACCGAGTGGGTGCTCTGCTCCGCTGGGAATTCATGATCTGA 175  
Db |||||  
404 GCACGACAGCCTTACCGAGTGGGTGCTCTGCTCCGCTGGGAATTCATGATCTGA 463  
QY 176 TAGGCAAGGTTTACACGACAGAACCCAGTGGCCACATGCGGCTGGGCAAAACATA 235  
Db |||||  
464 TAGGCAAGGTTTACACGACAGAACCCAGTGGCCACATGCGGCTGGGCAAAACATA 523  
QY 236 ATTACTGCGGAAATCCTGATGGGATGCAAGCCCTGTGTGACATGCTGAAGAACGCA 295  
Db |||||  
524 ATTACTGCGGAAATCCTGATGGGATGCAAGCCCTGTGTGACATGCTGAAGAACGCA 583  
QY 296 GGCTGACGTGGAGTACTGTGATGTCCTCTGCTCCACCTGCGGCTGAGACAGTACA 355  
Db |||||  
584 GGCTGACGTGGAGTACTGTGATGTCCTCTGCTCCACCTGCGGCTGAGACAGTACA 643  
QY 356 GCCAGCCTCAGTTTCGATCAAGAGGAGGCTCTTCCGCGACATCGCCTCCACCCCTGGC 415  
Db |||||  
644 GCCAGCCTCAGTTTCGATCAAGAGGAGGCTCTTCCGCGACATCGCCTCCACCCCTGGC 703  
QY 416 AGGCTGCATCTTTGCCAAGCAGAGAGTGGCCGAGAGGGTCTGTGCGGGGCA 475  
Db |||||  
704 AGGCTGCATCTTTGCCAAGCAGAGAGTGGCCGAGAGGGTCTGTGCGGGGCA 763  
QY 476 TACTCATCAGCTCTGCTGGATTCCTCTGCGCGCCACTGCTCCAGGAGAGTTTCGCG 535  
Db |||||  
764 TACTCATCAGCTCTGCTGGATTCCTCTGCGCGCCACTGCTCCAGGAGAGTTTCGCG 823  
QY 536 CCCACCACCTGACGGTGTCTTTGGGCAAGACATACCGGGTGTCTTCCGCGAGGAGGAGC 595  
Db |||||  
824 CCCACCACCTGACGGTGTCTTTGGGCAAGACATACCGGGTGTCTTCCGCGAGGAGGAGC 883  
QY 596 AGAAATTTGAGTCAAAAATACATTTCCATAGGAATTCATGATGACACTTACGACA 655  
Db |||||  
884 AGAAATTTGAGTCAAAAATACATTTCCATAGGAATTCATGATGACACTTACGACA 943  
QY 656 ATGACATTCGCTGCTGACGTGAATCGGATTCCTGCGCGCCACTGCTCCAGGAGAGCAGCG 715  
Db |||||  
944 ATGACATTCGCTGCTGACGTGAATCGGATTCCTGCGCGCCACTGCTCCAGGAGAGCAGCG 1003  
QY 716 TGGTCCGCACTGTGTGCTTCCCGCGGACCTGACGTGCGGACTGGACGAGTGTG 775  
Db |||||  
1004 TGGTCCGCACTGTGTGCTTCCCGCGGACCTGACGTGCGGACTGGACGAGTGTG 1063  
QY 776 AGCTCTCCGCTACGGCAAGCATGAGGCTGTGTCTCTTCTATTCGAGCGGCTGAAGG 835  
Db |||||  
1064 AGCTCTCCGCTACGGCAAGCATGAGGCTGTGTCTCTTCTATTCGAGCGGCTGAAGG 1123  
QY 836 AGGCTCATGTGACGTGTACCATCCAGCGCTGACATCAACATTTACTTAACAGAA 895  
Db |||||  
1124 AGGCTCATGTGACGTGTACCATCCAGCGCTGACATCAACATTTACTTAACAGAA 1183  
QY 896 CAGTCACCGACAACTGTGTGTGTGTGAGACACTCGGAGCGGGGGCCCGACGCAACT 955  
Db |||||  
1184 CAGTCACCGACAACTGTGTGTGTGTGAGACACTCGGAGCGGGGGCCCGACGCAACT 1243  
QY 956 TGCAGGACGCTGCCAGGGGAGTTCGGAGGGCCCTGTGTGTGTGATGATGCGCGCA 1015  
Db |||||  
1244 TGCAGGACGCTGCCAGGGGAGTTCGGAGGGCCCTGTGTGTGTGATGATGCGCGCA 1303  
QY 1016 TGACTTTGGTGGGACATCATGAGTGGGCTGTGGACAGAGGATGTCCCGGCTG 1075  
Db |||||  
1304 TGACTTTGGTGGGACATCATGAGTGGGCTGTGGACAGAGGATGTCCCGGCTG 1363  
QY 1076 TGTACAAAAGGTTTACCAACTACCTAGACTGGATTCGTGACAAACATGCGACCG 1128  
Db |||||

Db 1364 TGTACAAAAGGTTTACCAACTACCTAGACTGGATTCGTGACAAACATGCGACCG 1416

RESULT 2  
US-08-814-412-10  
; Sequence 10, Application US/08814412  
; Patent No. 6150141  
; GENERAL INFORMATION:  
; APPLICANT: Jarrell Ph.D., Kevin A.  
; TITLE OF INVENTION: Intron-Mediated Techniques and Reagents  
; NUMBER OF SEQUENCES: 46  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Choate, Hall & Stewart  
; STREET: 53 State Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/814,412  
; FILING DATE: 11-MAR-1997  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Jarrell Ph.D., Brenda H.  
; REGISTRATION NUMBER: 39,223  
; REFERENCE/DOCKET NUMBER: 0079571-0040  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617 248 5000  
; TELEFAX: 617 248 4000  
; INFORMATION FOR SEQ ID NO: 10:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1848 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; IMMEDIATE SOURCE:  
; CLONE: amplified t-PA clone  
US-08-814-412-10

Query Match 94.4%; Score 1065; DB 3; Length 1848;  
Best Local Similarity 99.5%; Pred. No. 8.6e-290;  
Matches 1068; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 56 CCCAGGCGCCTCTGAGGAAACAGTGTACTTTGGGAATGGGTACGCTTACCGTG 115  
Db |||||  
685 CCCCTGCCCTGCTCTGAGGAAACAGTGTACTTTGGGAATGGGTACGCTTACCGTG 744  
QY 116 GCACGACAGCCTTACACGACAGAACCCAGTGGCCACATGCGGCTGGGCAAAACATA 175  
Db |||||  
745 GCACGACAGCCTTACACGACAGAACCCAGTGGCCACATGCGGCTGGGCAAAACATA 804  
QY 176 TAGGCAAGGTTTACACGACAGAACCCAGTGGCCACATGCGGCTGGGCAAAACATA 235  
Db |||||  
805 TAGGCAAGGTTTACACGACAGAACCCAGTGGCCACATGCGGCTGGGCAAAACATA 864  
QY 236 ATTACTGCGGAAATCCTGATGGGATGCAAGCCCTGTGTGACATGCTGAAGAACGCA 295  
Db |||||  
865 ATTACTGCGGAAATCCTGATGGGATGCAAGCCCTGTGTGACATGCTGAAGAACGCA 924  
QY 296 GGCTGACGTGGAGTACTGTGATGTCCTCTGCTCCACCTGCGGCTGAGACAGTACA 355  
Db |||||  
925 GGCTGACGTGGAGTACTGTGATGTCCTCTGCTCCACCTGCGGCTGAGACAGTACA 984  
QY 356 GCCAGCCTCAGTTTCGATCAAGAGGAGGCTCTTCCGCGACATGCGCTCCACCCCTGGC 415  
Db |||||  
985 GCCAGCCTCAGTTTCGATCAAGAGGAGGCTCTTCCGCGACATGCGCTCCACCCCTGGC 1044



```

QY 416 AGGCTGCCATCTTTGCCAAGCAGAGAGTCCCGGAGAGCGGTTCTGTGCGGGGCA 475
Db 1045 AGGCTGCCATCTTTGCCAAGCAGAGAGTCCCGGAGAGCGGTTCTGTGCGGGGCA 1104
QY 476 TACTCATCAGCTCCCTGCTGGATTCTCTCTGCGGCCACTGCTTCCAGAGAGTTTCGC 535
Db 1105 TACTCATCAGCTCCCTGCTGGATTCTCTCTGCGGCCACTGCTTCCAGAGAGTTTCGC 1164
QY 536 CCCACCACCTGACGGTGATCTTTGGCAGAACATACCGGGTGCTCCCTGCGGAGGAGGAGC 595
Db 1165 CCCACCACCTGACGGTGATCTTTGGCAGAACATACCGGGTGCTCCCTGCGGAGGAGGAGC 1224
QY 596 AGAAATTTGAAGTCGAAAAATACATTGTCCATAGAAATTCGATGACACTTACGACA 655
Db 1225 AGAAATTTGAAGTCGAAAAATACATTGTCCATAGAAATTCGATGACACTTACGACA 1284
QY 656 ATGACATTGCGCTGCTGAGCTGAATCGGATTCGCTCCGCTGTGCCAGAGAGCAGCG 715
Db 1285 ATGACATTGCGCTGCTGAGCTGAATCGGATTCGCTCCGCTGTGCCAGAGAGCAGCG 1344
QY 716 TGTCTCGCACCTGTGTGCTTCCCGCGGGGACCTGCAGCTGCGGACTGGACGGAGTG 775
Db 1345 TGTCTCGCACCTGTGTGCTTCCCGCGGGGACCTGCAGCTGCGGACTGGACGGAGTG 1404
QY 776 AGCTCTCGGCTACGGCAAGCAGATGAGCGCTTGTCTCTTCTATTCGAGCGGCTGAAGG 835
Db 1405 AGCTCTCGGCTACGGCAAGCAGATGAGCGCTTGTCTCTTCTATTCGAGCGGCTGAAGG 1464
QY 836 AGGCTCATGTAGACTGTACCCATCCAGCCGCTGCACATCAACATTTACTTACAGAA 895
Db 1465 AGGCTCATGTAGACTGTACCCATCCAGCCGCTGCACATCAACATTTACTTACAGAA 1524
QY 896 CAGTCACCGCAACATGTGTGTGTGAGACACTCGGAGCGGGCGGCCAGGCAAACT 955
Db 1525 CAGTCACCGCAACATGTGTGTGTGAGACACTCGGAGCGGGCGGCCAGGCAAACT 1584
QY 956 TGCACAGCGCTGCCAGCGGCGATTCCGAGAGCGCCCTGGTGTCTGAACGATGCCGCA 1015
Db 1585 TGCACAGCGCTGCCAGCGGCGATTCCGAGAGCGCCCTGGTGTCTGAACGATGCCGCA 1644
QY 1016 TGACTTTGGTGGGCATCATCAGCTGGGCGCTGGGCTGTGGACAGAGATGTCGGG 1075
Db 1645 TGACTTTGGTGGGCATCATCAGCTGGGCGCTGGGCTGTGGACAGAGATGTCGGG 1704
QY 1076 TGTACAAAAGTTACCAACTACCTAGACTGATTCGTGACAAATGCGACG 1128
Db 1705 TGTACAAAAGTTACCAACTACCTAGACTGATTCGTGACAAATGCGACG 1757

```

## RESULT 3

```

US-08-811-949-42
; Sequence 42, Application US/08811949
; Patent No. 5840533
; GENERAL INFORMATION:
; APPLICANT: NIWA, MINEO
; APPLICANT: SAITO, YOSHIMASA
; APPLICANT: SASAKI, HITOSHI
; APPLICANT: HAYASHI, MASAKO
; APPLICANT: NOTANI, JOUJI
; APPLICANT: KOBAYASHI, MASAKAZU
; TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR
; NUMBER OF SEQUENCES: 67
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,
; ADDRESSEE: P.C.
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

```

```

; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/811,949
; FILING DATE: 05-MAR-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 18-966-0
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 42:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2101 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 25...1710
US-08-811-949-42

```

```

Query Match      94.4%; Score 1065; DB 2; Length 2101;
Best Local Similarity 99.5%; Pred. No. 9,1e-290;
Matches 1068; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```

```

QY 56 CCCAGGCGGCTCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTACGCTACCGTG 115
Db 638 CCCCTGCTGCTCTGAGGGAAACAGTGACTGCTACTTTGGGAATGGGTACGCTACCGTG 697
QY 116 GCACGACAGCTCACCGAGTCGGGTGCTCTCTCCCTCCGCTGGAATTCATGATCCTGA 175
Db 698 GCACGACAGCTCACCGAGTCGGGTGCTCTCTCCCTCCGCTGGAATTCATGATCCTGA 757
QY 176 TAGCAAGAGTTTACACAGCAGAACCCAGTGCCAGGCACTGGGCGCTGGGCAACATA 235
Db 758 TAGCAAGAGTTTACACAGCAGAACCCAGTGCCAGGCACTGGGCGCTGGGCAACATA 817
QY 236 ATTACTCCGGAATCTGTATGGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCA 295
Db 818 ATTACTCCGGAATCTGTATGGGGATGCCAAGCCCTGGTGCCACGTGCTGAAGAACCGCA 877
QY 296 GGCTGAGTGGGAGTACTGTGATGTGCTCTCTCCCTCCAGCTGGGCTGAGACAGTACA 355
Db 878 GGCTGAGTGGGAGTACTGTGATGTGCTCTCTCCCTCCAGCTGGGCTGAGACAGTACA 937
QY 356 GCCAGCTCTCAGTTTCGCATCAAGAGGAGGCTCTTTCGCCGACATCGCTCCCACTCCCTGGC 415
Db 938 GCCAGCTCTCAGTTTCGCATCAAGAGGAGGCTCTTTCGCCGACATCGCTCCCACTCCCTGGC 997
QY 416 AGGTGCGCATCTTTGCGCAAGCAGAGAGTGCCTCCGAGAGCGGTTCTGTGCGGGGCA 475
Db 998 AGGTGCGCATCTTTGCGCAAGCAGAGAGTGCCTCCGAGAGCGGTTCTGTGCGGGGCA 1057
QY 476 TACTCATCAGCTCCTGCTGGATTCTCTCTGCGGCCACTGCTTCCAGAGAGGTTTCGC 535
Db 1058 TACTCATCAGCTCCTGCTGGATTCTCTCTGCGGCCACTGCTTCCAGAGAGGTTTCGC 1117
QY 536 CCCACCACCTGACGGTGATCTTTGGGAGAACATACCGGGTGCTCCCTGCGGAGGAGGAGC 595
Db 1118 CCCACCACCTGACGGTGATCTTTGGGAGAACATACCGGGTGCTCCCTGCGGAGGAGGAGC 1177
QY 596 AGAAATTTGAAGTCGAAAAATACATTGTCCATAGAAATTCGATGACACTTACGACA 655
Db 1178 AGAAATTTGAAGTCGAAAAATACATTGTCCATAGAAATTCGATGACACTTACGACA 1237
QY 656 ATGACATTGCGCTGCTGAGCTGAATCGGATTCGCTCCGCTGTGCCAGAGAGCAGCG 715
Db 1238 ATGACATTGCGCTGCTGAGCTGAATCGGATTCGCTCCGCTGTGCCAGAGAGCAGCG 1297

```

QY 716 TGGTCCGCACTGTGTGCTTCCCGGGGACCTGACGTGCGGACCTGGAGGAGTGTG 775  
Db 1298 TGGTCCGCACTGTGTGCTTCCCGGGGACCTGACGTGCGGACCTGGAGGAGTGTG 1357  
QY 776 AGCTCTCCGGCTACCGCAAGCATGAGGCTTGTCTCTTTCTATTTCGGAGCGGCTGAAGG 835  
Db 1358 AGCTCTCCGGCTACCGCAAGCATGAGGCTTGTCTCTTTCTATTTCGGAGCGGCTGAAGG 1417  
QY 836 AGGCTCATGTGACATGTATCCCATCCAGCGCTGCACATCAACATTTACTTAACAGAA 895  
Db 1418 AGGCTCATGTGACATGTATCCCATCCAGCGCTGCACATCAACATTTACTTAACAGAA 1477  
QY 896 CAGTCACCGCAACATGTGTGTGTGAGACACTCGGAGCGCGGGCCCGAGGCAAACT 955  
Db 1478 CAGTCACCGCAACATGTGTGTGTGAGACACTCGGAGCGCGGGCCCGAGGCAAACT 1537  
QY 956 TGCACGAGCGCTGCCAGGCGGATTCGGAGGCGCCCTGGTGTCTGAACGATGCGCGCA 1015  
Db 1538 TGCACGAGCGCTGCCAGGCGGATTCGGAGGCGCCCTGGTGTCTGAACGATGCGCGCA 1597  
QY 1016 TGACTTTGCTGGGCATCATCAGCTGGGCGCTGGGCTGTGGACAGAGGATGTCCCGGGTG 1075  
Db 1598 TGACTTTGCTGGGCATCATCAGCTGGGCGCTGGGCTGTGGACAGAGGATGTCCCGGGTG 1657  
QY 1076 TGTACAAAAGGTTACCAACTACCTAGACTGGAATTCGTGACAAACATGCGACCG 1128  
Db 1658 TGTACAAAAGGTTACCAACTACCTAGACTGGAATTCGTGACAAACATGCGACCG 1710

## RESULT 4

US-08-119-512-3  
; Sequence 3, Application US/08119512  
; Patent No. 5498531  
; GENERAL INFORMATION:  
; APPLICANT: Jarrell, Kevin A.  
; TITLE OF INVENTION: INTRON MEDIATED RECOMBINANT TECHNIQUES  
; TITLE OF INVENTION: AND REAGENTS  
; NUMBER OF SEQUENCES: 3  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: LAHIVE & COCKFIELD  
; STREET: 60 State Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: ASCII (text)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/119,512  
; FILING DATE: 10-SEP-1993  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Vincent, Matthew P.  
; REGISTRATION NUMBER: 36,709  
; REFERENCE/DOCKET NUMBER: HUI-008  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 227-7400  
; TELEFAX: (617) 227-5941  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2162 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: both  
; MOLECULE TYPE: other nucleic acid  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: 82..334  
; OTHER INFORMATION: /product= "Signal Sequence and  
; OTHER INFORMATION: Finger-like domain"

; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: 335..447  
; OTHER INFORMATION: /product= "EGF-like domain"  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: 448..714  
; OTHER INFORMATION: /product= "Kringle-1 domain"  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: 715..972  
; OTHER INFORMATION: /product= "Kringle-2 domain"  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: 973..2162  
; OTHER INFORMATION: /product= "Catalytic domain"  
; US-08-119-512-3  
Query Match 94.4%; Score 1065; DB 1; Length 2162;  
Best Local Similarity 99.5%; Pred. No. 9.2e-290;  
Matches 1068; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
QY 56 CCCAGGCGGCTCTGAGGGAAACAGTCACTGTCTACTTTTGGGAATGGGTCACGCTACCGTG 115  
Db 695 CCCCTGCTGCTCTGAGGGAAACAGTCACTGTCTACTTTTGGGAATGGGTCACGCTACCGTG 754  
QY 116 GCACGACAGCTCACCGAGTCGGGTGCTCTCTCTCCCTCCGCTGGAATTCATGATCCTGA 175  
Db 755 GCACGACAGCTCACCGAGTCGGGTGCTCTCTCTCCCTCCGCTGGAATTCATGATCCTGA 814  
QY 176 TAGCAAGGTTTACACAGCACAGAAACCCAGTGCACAGGACACTGGGGCTTGGGCAACATA 235  
Db 815 TAGCAAGGTTTACACAGCACAGNAACCCAGTGCACAGGACACTGGGGCTTGGGCAACATA 874  
QY 236 ATTACTCCGGAATCTGATGGGATCCAAAGCCCTGGTGCACAGTCTGAAGAACCGCA 295  
Db 875 ATTACTCCGGAATCTGATGGGATCCAAAGCCCTGGTGCACAGTCTGAAGAACCGCA 934  
QY 296 GGCTGAGTGGGAGTACTGTGTCCTCTCTGCTCCACTGCGGCTGAGACAGTACA 355  
Db 935 GGCTGAGTGGGAGTACTGTGTCCTCTCTGCTCCACTGCGGCTGAGACAGTACA 994  
QY 356 GCCAGCCTCAGTTTCGATCAAAAGGAGGGCTCTTTCGCGGACATCGCCTCCACCCCTGGC 415  
Db 995 GCCAGCCTCAGTTTCGATCAAAAGGAGGGCTCTTTCGCGGACATCGCCTCCACCCCTGGC 1054  
QY 416 AGGCTGCCATCTTTGCCAAGCACAGAGGTGCGCCGAGAGCGGTTCTGTGCGGGGCA 475  
Db 1055 AGGCTGCCATCTTTGCCAAGCACAGAGGTGCGCCGAGAGCGGTTCTGTGCGGGGCA 1114  
QY 476 TACTCATCAGCTCTGCTGGATTTCTCTCTCGCGCCACTGCTTCCAGGAGGTTTCCGC 535  
Db 1115 TACTCATCAGCTCTGCTGGATTTCTCTCTCGCGCCACTGCTTCCAGGAGGTTTCCGC 1174  
QY 536 CCACACACCTGACGGTGATCTTGGGAGAAACATACCGGGTGGTCCCTGGCGAGGAGGAGC 595  
Db 1175 CCACACACCTGACGGTGATCTTGGGAGAAACATACCGGGTGGTCCCTGGCGAGGAGGAGC 1234  
QY 596 AGAAATTTGAAGTCGAAAAATACATTGTCCATAGGAATTCGATGATGACATTTACGACA 655  
Db 1235 AGAAATTTGAAGTCGAAAAATACATTGTCCATAGGAATTCGATGATGACATTTACGACA 1294  
QY 656 ATGACATTTGGCTGCTGCAGCTGAATTCGATTTGCTCCGCTGTCGCCAGGAGGAGCG 715  
Db 1295 ATGACATTTGGCTGCTGCAGCTGAATTCGATTTGCTCCGCTGTCGCCAGGAGGAGCG 1354  
QY 716 TGGTCCGCACTGTGTGCTTCCCGCGGACCTGACAGCTGCGGAGTGGACCGAGGTGTG 775  
Db 1355 TGGTCCGCACTGTGTGCTTCCCGCGGACCTGACAGCTGCGGAGTGGACCGAGGTGTG 1414  
QY 776 AGCTCTCCGGCTACCGCAAGCATGAGGCTTGTCTCTTTCTATTTCGGAGCGGCTGAAGG 835  
Db 1415 AGCTCTCCGGCTACCGCAAGCATGAGGCTTGTCTCTTTCTATTTCGGAGCGGCTGAAGG 1474

QY 836 AGGCTCATGTGAGTGTACCATCCAGCGGTGCACATCAACAATTTACTTAACGAA 895  
Db 1475 AGGCTCATGTGAGTGTACCATCCAGCGGTGCACATCAACAATTTACTTAACGAA 1534  
QY 896 CAGTCACCGACACATGCTGTGTGTGAGACACTCGGAGCGGGGGCCCGAGGCAAACT 955  
Db 1535 CAGTCACCGACACATGCTGTGTGTGAGACACTCGGAGCGGGGGCCCGAGGCAAACT 1594  
QY 956 TGCACGACGCTGCCAGGCGGATTCGGGAGGCGCCCTGTGTGTCTGAACGATGCGCGCA 1015  
Db 1595 TCACGACGCTGCCAGGCGGATTCGGGAGGCGCCCTGTGTGTCTGAACGATGCGCGCA 1654  
QY 1016 TGACTTTGGTGGGACATCATAGCTGGGCGCTGGGCTGTGGACAGAGGATGTCCGGGTG 1075  
Db 1655 TGACTTTGGTGGGACATCATAGCTGGGCGCTGGGCTGTGGACAGAGGATGTCCGGGTG 1714  
QY 1076 TGTACACAAAGTTACCAACTACCTAGACTGGATTCGTGTGACAACTGCGACCG 1128  
Db 1715 TGTACACAAAGTTACCAACTACCTAGACTGGATTCGTGTGACAACTGCGACCG 1767

RESULT 5

US-08-488-015B-3  
; Sequence 3, Application US/08488015B  
; Patent No. 5780272

GENERAL INFORMATION:

; APPLICANT: Jarrell, Kevin A.

; TITLE OF INVENTION: INTRON-MEDIATED RECOMBINANT TECHNIQUES

; TITLE OF INVENTION: AND REAGENTS

; NUMBER OF SEQUENCES: 29

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Foley, Hoag & Eliot

; STREET: One Post Office Square

; CITY: Boston

; STATE: MA

; COUNTRY: USA

; ZIP: 02109

COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: ASCII (text)

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/488,015B

; FILING DATE: 07-JUN-1995

; ATTORNEY/AGENT INFORMATION:

; NAME: Vincent, Matthew P.

; REGISTRATION NUMBER: 36,709

; REFERENCE/DOCKET NUMBER: HUV-008.02

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (617) 832-1000

; TELEFAX: (617) 832-7000

; INFORMATION FOR SEQ ID NO: 3:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 2162 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: double

; TOPOLOGY: both

; MOLECULE TYPE: other nucleic acid

; FEATURE:

; NAME/KEY: misc feature

; LOCATION: 82..334

; OTHER INFORMATION: /product= "Signal Sequence and

; Patent No. 5780272

; OTHER INFORMATION: Finger-like domain"

; FEATURE:

; NAME/KEY: misc feature

; LOCATION: 335..447

; OTHER INFORMATION: /product= "EGF-like domain"

; FEATURE:

; NAME/KEY: misc feature

; LOCATION: 448..714

; OTHER INFORMATION: /product= "Kringle-1 domain"

; FEATURE:

; NAME/KEY: misc feature

; LOCATION: 715..972

; OTHER INFORMATION: /product= "Kringle-2 domain"

; FEATURE:

; NAME/KEY: misc feature

; LOCATION: 973..2162

; OTHER INFORMATION: /product= "Catalytic domain"

; US-08-488-015B-3

Query Match 94.4%; Score 1065; DB 1; Length 2162;

Best Local Similarity 99.5%; Pred. No. 9.2e-290;

Matches 1069; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 56 CCACAGGCGGCTCTGAGGGAAACAGTGAAGTCTTGGCAATGGGTGAGCTTACCGTG 115

Db 695 CCCTGCTGCTCTGAGGGAAACAGTGAAGTCTTGGCAATGGGTGAGCTTACCGTG 754

QY 116 GCAGCGACAGCTTACCGAGTCCGGTGCCTCTGCTCCCTCCCGTGGAAATCCATGATCTGA 175

Db 755 GCAGCGACAGCTTACCGAGTCCGGTGCCTCTGCTCCCTCCCGTGGAAATCCATGATCTGA 814

QY 176 TAGCAAGGTTTACACAGCACAGAAACCCAGTCCAGGCACTGGGCTTGGGCAAAACATA 235

Db 815 TAGCAAGGTTTACACAGCACAGAAACCCAGTCCAGGCACTGGGCTTGGGCAAAACATA 874

QY 236 ATTACTGCGGAAATCTGATGGGATGCAAGCCCTGTCGCCAGTCTGCAAGAACCGCA 295

Db 875 ATTACTGCGGAAATCTGATGGGATGCAAGCCCTGTCGCCAGTCTGCAAGAACCGCA 934

QY 296 GGCTGAGTGGGAGTACTGTGTGCTCTCTGCTCCACTTGGCGCTGAGACAGTACA 355

Db 935 GGCTGAGTGGGAGTACTGTGTGCTCTCTGCTCCACTTGGCGCTGAGACAGTACA 994

QY 356 GCACAGCTCAGTTTCGATCAAAAGAGGAGTCTTTCGCCGAGATCGCTCCACCCCTGGC 415

Db 995 GCACAGCTCAGTTTCGATCAAAAGAGGAGTCTTTCGCCGAGATCGCTCCACCCCTGGC 1054

QY 416 AGGCTGCCATCTTTCGCAAGCAGAGGAGTTCGCCGAGAGCGGTCTCTGTCGGGGGCA 475

Db 1055 AGGCTGCCATCTTTCGCAAGCAGAGGAGTTCGCCGAGAGCGGTCTCTGTCGGGGGCA 1114

QY 476 TACTCATCAGTCTCTGCTGGATTTCTCTGTCGCCGCCACTGCTTCCAGGAGAGGTTTCCGC 535

Db 1115 TACTCATCAGTCTCTGCTGGATTTCTCTGTCGCCGCCACTGCTTCCAGGAGAGGTTTCCGC 1174

QY 536 CCACACCTGACCGTGTATCTTGGGAGAAACATACCGGGTGGTCCCTGGCGAGGAGGC 595

Db 1175 CCACACCTGACCGTGTATCTTGGGAGAAACATACCGGGTGGTCCCTGGCGAGGAGGC 1234

QY 596 AGAAATTTGAAGTCGAAATAATACATTCCTCAATAGGAATTCGATGATGACATTTACCACA 655

Db 1235 AGAAATTTGAAGTCGAAATAATACATTCCTCAATAGGAATTCGATGATGACATTTACCACA 1294

QY 656 ATGACATTTGGCTGCTGCAGCTGAAATCGGATTCGTCGCCGCTGTGCCAGGAGAGCG 715

Db 1295 ATGACATTTGGCTGCTGCAGCTGAAATCGGATTCGTCGCCGCTGTGCCAGGAGAGCG 1354

QY 716 TGGTCCGCACTGTGTGCTTCCCGCGGAGACCTGTCAGCTGCCGGAATGACGAGGTGTG 775

Db 1355 TGGTCCGCACTGTGTGCTTCCCGCGGAGACCTGTCAGCTGCCGGAATGACGAGGTGTG 1414

QY 776 AGCTTCGGCTAGCGCAAGCATGAGGCTTGTCTCTCTTCTATTTCGAGCGGCTGAAGG 835

Db 1415 AGCTTCGGCTAGCGCAAGCATGAGGCTTGTCTCTCTTCTATTTCGAGCGGCTGAAGG 1474

QY 836 AGGCTCATGTGAGTGTACCATCCAGCGGCTGCACATCAACAATTTACTTAACAGAA 895

Db 1475 AGGCTCATGTGAGTGTACCATCCAGCGGCTGCACATCAACAATTTACTTAACAGAA 1534

QY 896 CAGTCACCGACACATGCTGTGTGCTGAGACACTCGGAGCGGGGGCCCGAGGCAAACT 955

Db 955 CAGTCACCGACACATGCTGTGTGCTGAGACACTCGGAGCGGGGGCCCGAGGCAAACT 955

Db 1535 CAGTCACCGACAAACATGCTGTGTGTGGAGACACATCGGAGCGCGCGGCCCAAGCAAACT 1594  
QY 956 TGCACGACGCTCGCCAGGCGCAATTCGGAGAGGCCCTCGTGTGTCTGAACGATGCGCGCA 1015  
Db 1595 TGCACGACGCTCGCCAGGCGCAATTCGGAGAGGCCCTCGTGTGTCTGAACGATGCGCGCA 1654  
QY 1016 TGACTTTGGTGGGCAATCATCAGCTGGGGCTTGGGCTGTGGAGCAGAAAGATGTCGGGGTG 1075  
Db 1655 TGACTTTGGTGGGCAATCATCAGCTGGGGCTTGGGCTGTGGAGCAGAAAGATGTCGGGGTG 1714  
QY 1076 TGTACACAAAGGTTACCAACTACCTAGACTGATTCGTGACAAATGCGACG 1128  
Db 1715 TGTACACAAAGGTTACCAACTACCTAGACTGATTCGTGACAAATGCGACG 1767

## RESULT 6

US-08-488-015B-25.  
; Sequence 25, Application US/08488015B  
; Patent No. 5780272  
; GENERAL INFORMATION:  
; APPLICANT: Jarrell, Kevin A.  
; TITLE OF INVENTION: INTRON-MEDIATED RECOMBINANT TECHNIQUES  
; TITLE OF INVENTION: AND REAGENTS  
; NUMBER OF SEQUENCES: 29  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Foley, Hoag & Eliot  
; STREET: One Post Office Square  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: ASCII (text)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/488,015B  
; FILING DATE: 07-JUN-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Vincent, Matthew P.  
; REGISTRATION NUMBER: 36,709  
; REFERENCE/DOCKET NUMBER: HUV-008.02  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 832-1000  
; TELEFAX: (617) 832-7000  
; INFORMATION FOR SEQ ID NO: 25:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2162 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
US-08-488-015B-25

Query Match 94.4%; Score 1065; DB 1; Length 2162;

Best Local Similarity 99.5%; Pred. No. 9, 2e-290;

Matches 1068; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 56 CCCAGGGCGGCTCTGAGGGAACAGTGTACTTGGGAATGGGTGACGCTACCGTG 115  
Db 695 CCCCTGCTGCTCTGAGGGAACAGTGTACTTGGGAATGGGTGACGCTACCGTG 754  
QY 116 GCAGCAGAGCTCACCGAGTCGGGTGCTCTCGCTCCCGTGGAAATCCATGATCTGA 175  
Db 755 GCAGCAGAGCTCACCGAGTCGGGTGCTCTCGCTCCCGTGGAAATCCATGATCTGA 814  
QY 176 TAGGCAAGGTTTACACAGCAGACAGAACCCAGTCGCCAGGCACTGGGGCTGGGCAACATA 235  
Db 815 TAGGCAAGGTTTACACAGCAGACAGAACCCAGTCGCCAGGCACTGGGGCTGGGCAACATA 874  
QY 236 ATTACTGCGGAATCCTGATGGGATGCCAAGCCCTGGTGCACATGCTGTGAAGAACCGCA 295

Db 875 ATTACTGCGGAATCCTGATGGGATGCCAAGCCCTGGTGCACATGCTGTGAAGAACCGCA 934  
QY 296 GGCTGAGCTGGGAGTACTGTGATGTGCTGCTCCCTGCTCCACTGCGGGCTGAGACAGTACA 355  
Db 935 GGCTGAGCTGGGAGTACTGTGATGTGCTGCTCCCTGCTCCACTGCGGGCTGAGACAGTACA 994  
QY 356 GCCAGCCTCAGTTTTCGATCAAAAGGAGGGCTCTTTCGCCGACATCGCCTCCACACCCCTGGC 415  
Db 995 GCCAGCCTCAGTTTTCGATCAAAAGGAGGGCTCTTTCGCCGACATCGCCTCCACACCCCTGGC 1054  
QY 416 AGGCTGCATCTTTTGCACAAGCAGAGAGTTCGCCCGGAGAGCGGTTCTGTGCGGGGCGCA 475  
Db 1055 AGGCTGCATCTTTTGCACAAGCAGAGAGTTCGCCCGGAGAGCGGTTCTGTGCGGGGCGCA 1114  
QY 476 TACTCATCAGCTCCTGCTGATTCCTCTCGCGGCCACTGCTTCCAGAGAGGTTTCCGC 535  
Db 1115 TACTCATCAGCTCCTGCTGATTCCTCTCGCGGCCACTGCTTCCAGAGAGGTTTCCGC 1174  
QY 536 CCCACCACTCAGCGGTGATCTTGGGAGAAACATACCGGGTGGTCCCTGGCGAGAGGAGC 595  
Db 1175 CCCACCACTCAGCGGTGATCTTGGGAGAAACATACCGGGTGGTCCCTGGCGAGAGGAGC 1234  
QY 596 AGAAATTTGAAGTCGAAAAATATCTGTCCATAGGAATTCGATGATGACACTTACGACA 655  
Db 1235 AGAAATTTGAAGTCGAAAAATATCTGTCCATAGGAATTCGATGATGACACTTACGACA 1294  
QY 656 ATGACATTTGCGCTGCTGACGCTGAAATCGGATTCGTCCTCCCTGTCGCCAGGAGGAGCG 715  
Db 1295 ATGACATTTGCGCTGCTGACGCTGAAATCGGATTCGTCCTCCCTGTCGCCAGGAGGAGCG 1354  
QY 716 TGGTCCGCACTGTGTCTTCCCGCGGAGCACTGCACTGCGGCACTGGACGAGGTGTG 775  
Db 1355 TGGTCCGCACTGTGTCTTCCCGCGGAGCACTGCACTGCGGCACTGGACGAGGTGTG 1414  
QY 776 AGCTCTCCGCTAGCGCAAGCATGAGGCTTGTCTCTCTTCTATTTCGAGCGGCTGAAG 835  
Db 1415 AGCTCTCCGCTAGCGCAAGCATGAGGCTTGTCTCTCTTCTATTTCGAGCGGCTGAAG 1474  
QY 836 AGGCTCATGTGCACTGTATCCCATCCAGCGCTGCACTCACAACATTTTAAACAGAA 895  
Db 1475 AGGCTCATGTGCACTGTATCCCATCCAGCGCTGCACTCACAACATTTTAAACAGAA 1534  
QY 896 CAGTCACCGCAACATGCTGTGTGGAGACACTCGGAGCGGGCGGCCAGGCAAACT 955  
Db 1535 CAGTCACCGCAACATGCTGTGTGGAGACACTCGGAGCGGGCGGCCAGGCAAACT 1594  
QY 956 TGACAGACGCTCGCAGGGCGATTCGGAGAGGCCCTGCTGTGTCTGAAAGATGCGCGCA 1015  
Db 1595 TGACAGACGCTCGCAGGGCGATTCGGAGAGGCCCTGCTGTGTCTGAAAGATGCGCGCA 1654  
QY 1016 TGACTTTGGTGGGCATCATCAGCTGGGGCTGGGCTGTGGACAGAGGATGTCGGGGTG 1075  
Db 1655 TGACTTTGGTGGGCATCATCAGCTGGGGCTGGGCTGTGGACAGAGGATGTCGGGGTG 1714  
QY 1076 TGTACACAAAGGTTTACCAACTACCTAGACTGATTCGTGACAAATGCGACG 1128  
Db 1715 TGTACACAAAGGTTTACCAACTACCTAGACTGATTCGTGACAAATGCGACG 1767

## RESULT 7

US-08-811-949-48  
; Sequence 48, Application US/08811949  
; Patent No. 5840533  
; GENERAL INFORMATION:  
; APPLICANT: NIWA, MINEO  
; APPLICANT: SAITO, YOSHIMASA  
; APPLICANT: SASAKI, HITOSHI  
; APPLICANT: HAYASHI, MASAKO  
; APPLICANT: NOTANI, JOUJI  
; APPLICANT: KOBAYASHI, MASAKAZU  
; TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR  
; NUMBER OF SEQUENCES: 67  
; CORRESPONDENCE ADDRESS:

ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
ADDRESS: P.C.  
STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400  
CITY: ARLINGTON  
STATE: VA  
COUNTRY: USA  
ZIP: 22202  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/811,949  
FILING DATE: 05-MAR-1997  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: OBLON, NORMAN F.  
REGISTRATION NUMBER: 24,618  
REFERENCE/DOCKET NUMBER: 18-966-0  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-413-3000  
TELEFAX: 703-413-2220  
INFORMATION FOR SEQ ID NO.: 48:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1314 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: circular  
MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..1311  
US-08-811-949-48

Query Match 94.3%; Score 1063.4; DB 2; Length 1314;  
Best Local Similarity 99.4%; Pred. No. 2.1e-289;  
Matches 1067; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 56 CCCAGCGGCTCTGAGGAAACAGTACTGCTCTTTGGGAATGGGTACGCTACCGTG 115  
DB 239 CCCCTGCTCTGAGGAAACAGTACTGCTCTTTGGGAATGGGTACGCTACCGTG 298  
QY 116 GCACGACAGCTCAACGAGTGGGTGCTCTGCTCCCGTGGGAATCCATGATCTGA 175  
DB 299 GCAGCAGAGCTCAACGAGTGGGTGCTCTGCTCCCGTGGGAATCCATGATCTGA 358  
QY 176 TAGGCAAGTTTACACAGACAGAACCCAGTCCCGAGGCACTGGGCTGGGCAACATA 235  
DB 359 TAGGCAAGTTTACACAGACAGAACCCAGTCCCGAGGCACTGGGCTGGGCAACATA 418  
QY 236 ATTACTGCGGAATCTGATGGGATGCAAGCCCTGGTGCACGTCTGAAGAACGCA 295  
DB 419 ATTACTGCGGAATCTGATGGGATGCAAGCCCTGGTGCACGTCTGAAGAACGCA 478  
QY 296 GGTGACGTGGGAGTACTGATGTCCTCTGCTCCACCTGGGCTGGGCAAGATACA 355  
DB 479 GGTGACGTGGGAGTACTGATGTCCTCTGCTCCACCTGGGCTGGGCAAGATACA 538  
QY 356 GCCAGCTCAGTTTCGATCAAGAGGCTCTTCGCCGACATGCTCCACCCCTGGC 415  
DB 539 GCCAGCTCAGTTTCGATCAAGAGGCTCTTCGCCGACATGCTCCACCCCTGGC 598  
QY 416 AGGTGCGCATCTTTGCCAAGCAGAGAGTCCCGGAGAGCGGTTCTGTCGGGGCA 475  
DB 599 AGGTGCGCATCTTTGCCAAGCAGAGAGTCCCGGAGAGCGGTTCTGTCGGGGCA 658  
QY 476 TACTCATCAGTCTCTGCTGGATTCTCTCTGCGGCCACTGCTTCAGGAGAGTTTCGC 535  
DB 659 TACTCATCAGTCTCTGCTGGATTCTCTCTGCGGCCACTGCTTCAGGAGAGTTTCGC 718  
QY 536 CCCACCACTGACGGTGTATCTTTGGGCAACATACCGGGTGGTCCCTGGGAGGAGC 595

DB 719 CCCACCACTGACGGTGTATCTTTGGGCAAGACATACCGGGTGGTCCCTGGGAGGAGC 778  
QY 596 AGAAATTTGAAGTCGAAAAATACATTTGTCATAGGAATTCGATGACACACTTTACGACA 655  
DB 779 AGAAATTTGAAGTCGAAAAATACATTTGTCATAGGAATTCGATGACACACTTTACGACA 838  
QY 656 ATGACATTTGCGCTGCTGACGCTGAAATCGGATTCGCTCCCGTGTGCCAGAGAGCAGCG 715  
DB 839 ATGACATTTGCGCTGCTGACGCTGAAATCGGATTCGCTCCCGTGTGCCAGAGAGCAGCG 898  
QY 716 TGGTCCGCACTGTGTGCTTTCCCGCGGACCTGCGAGCTGCGGACTTGGACGAGTGTG 775  
DB 899 TGGTCCGCACTGTGTGCTTTCCCGCGGACCTGCGAGCTGCGGACTTGGACGAGTGTG 958  
QY 776 AGCTCTCCGGCTACGGCAAGCATGAGGCTTTGTCTCTCTTTCTATTCGAGCGGCTGAAG 835  
DB 959 AGCTCTCCGGCTACGGCAAGCATGAGGCTTTGTCTCTCTTTCTATTCGAGCGGCTGAAG 1018  
QY 836 AGGCTCATGTGACACTGTATCCATCCAGCGCTGCACATACACAACTTTACTTAAACAGAA 895  
DB 1019 AGGCTCATGTGACACTGTATCCATCCAGCGCTGCACATCACAACTTTACTTAAACAGAA 1078  
QY 896 CAGTCACCGCAACATCTGTGTGCTGAGACACTCGGAGCGGCGGCCCCAGGCAAACT 955  
DB 1079 CAGTCACCGCAACATCTGTGTGCTGAGACACTCGGAGCGGCGGCCCCAGGCAAACT 1138  
QY 956 TGACACGACGCTGCGAGGCGAATTCGGGAGCGCCCTGGTGTGCTGAACGATGCGCGCA 1015  
DB 1139 TGACACGACGCTGCGAGGCGAATTCGGGAGCGCCCTGGTGTGCTGAACGATGCGCGCA 1198  
QY 1016 TGACTTTGGTGGGCACTCATCAGTGGGCTGTGGGCTGTGGACAGAGATGTCGCGGTG 1075  
DB 1199 TCACCTTTGGTGGGCACTCATCAGTGGGCTGTGGGCTGTGGACAGAGATGTCGCGGTG 1258  
QY 1076 TGTACAAAGGTTTACCACTACCTAGACTGGATTCGTGACAACTGCGACG 1128  
DB 1259 TGTACAAAGGTTTACCACTACCTAGACTGGATTCGTGACAACTGCGACG 1311

RESULT 8  
5200340-1  
; Patent No. 5200340  
; APPLICANT: FOSTER, DONALD C.; MULVIHILL, EILEEN R.; O'HARA,  
; PATRICK J.; PINGEL, KURT; YOSHITAKE, SHINJI  
; TITLE OF INVENTION: THROMBIN-ACTIVATED TISSUE PLASMINOGEN  
; ACTIVATORS  
; NUMBER OF SEQUENCES: 34  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/53,412  
; FILING DATE: 22-MAY-1987  
; SEQ ID NO: 1:  
; LENGTH: 1738  
5200340-1

Query Match 94.3%; Score 1063.4; DB 6; Length 1738;  
Best Local Similarity 99.4%; Pred. No. 2.4e-289;  
Matches 1067; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 56 CCCAGCGGCTCTGAGGAAACAGTACTGCTCTTTGGGAATGGGTACGCTACCGTG 115  
DB 629 CCCCTGCTCTGAGGAAACAGTACTGCTCTTTGGGAATGGGTACGCTACCGTG 688  
QY 116 GCACGACAGCTCAGGAGTCCCGTCCCTGCTCCCGTGGAAATTCATGATCTGA 175  
DB 689 GCACGACAGCTCAGGAGTCCCGTCCCTGCTCCCGTGGAAATTCATGATCTGA 748  
QY 176 TAGGCAAGTTTACACAGCAGAACCCAGTCCCGAGGCACTGGGCTGGGCAACATA 235  
DB 749 TAGGCAAGTTTACACAGCAGAACCCAGTCCCGAGGCACTGGGCTGGGCAACATA 808  
QY 236 ATTACTGCGGAATCTGATGGGATGCGAAGCCCTGGTGCACGTGCTGAGAACGCA 295  
DB 809 ATTACTGCGGAATCTGATGGGATGCGAAGCCCTGGTGCACGTGCTGAGAACGCA 868



QY 656 ATGACATTCGCTGCTGACGCTGAATCGGATTCGTCGGCTGTGCTCCGACGAGACAGCG 715  
Db 1304 ATGACATTCGCTGCTGACGCTGAATCGGATTCGTCGGCTGTGCTCCGACGAGACAGCG 1363  
QY 716 TGGTCCGACATGCTGCTGCTCCCGGGGACCTGACGCTGCGGACCTGGACGAGTGTG 775  
Db 1364 TGGTCCGACATGCTGCTGCTCCCGGGGACCTGACGCTGCGGACCTGGACGAGTGTG 1423  
QY 776 AGCTCTCGGCTPACGGCAAGCATAGGCTGTGCTCTCTTTCTATTCGGAGCGCTGAAGG 835  
Db 1424 AGCTCTCGGCTPACGGCAAGCATAGGCTGTGCTCTCTTTCTATTCGGAGCGCTGAAGG 1483  
QY 836 AGGCTCATGCTGACATGCTACCATCCAGCGCTGACATCAACAATTTACTTAACAGAA 895  
Db 1484 AGGCTCATGCTGACATGCTACCATCCAGCGCTGACATCAACAATTTACTTAACAGAA 1543  
QY 896 CAGTCACCGACAATGCTGTGCTGGAGACATCGGAGCGCGGGCCCGAGGCAAACT 955  
Db 1544 CAGTCACCGACAATGCTGTGCTGGAGACATCGGAGCGCGGGCCCGAGGCAAACT 1603  
QY 956 TGCAGACGCTGCCAGGCGCATTCGGAGGCGCCCTGCTGTGCTGAACGATGCGCGCA 1015  
Db 1604 TGCAGACGCTGCCAGGCGCATTCGGAGGCGCCCTGCTGTGCTGAACGATGCGCGCA 1663  
QY 1016 TGACTTTGGTGGGATCATCAGCTGGGCGCTGGGCTGTGGACAGAGATGTCGGGTG 1075  
Db 1664 TGACTTTGGTGGGATCATCAGCTGGGCGCTGGGCTGTGGACAGAGATGTCGGGTG 1723  
QY 1076 TGTACAAAAGTTACCAACTACCTAGACTGATTCGTCGACAACTGCGACG 1128  
Db 1724 TGTACAAAAGTTACCAACTACCTAGACTGATTCGTCGACAACTGCGACG 1776

RESULT 10  
5344773-1  
;Patent No. 5344773  
;APPLICANT: WEI, CHA-MER,HSIUNG, NANCY;REDDY, VERMURI B.;  
;LEMONTT, JEFFREY F.;DACKOWSKI, WILLIAM;DOUGLAS, RICHARD;  
;COLE, EDWARD S.;PURCELL JR., RICHARD D.;LAU, DAVID TAI-YUI  
; TITLE OF INVENTION: HUMAN UTERINE TISSUE PLASMINOGEN  
;ACTIVATOR PRODUCED BY RECOMBIANT DNA  
; NUMBER OF SEQUENCES: 6  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/782.686  
; FILING DATE: 01-OCT-1985  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 656.770  
; FILING DATE: 01-OCT-1984  
; SEQ ID NO:1:  
; LENGTH: 2457  
5344773-1  
Query Match 94.3%; Score 1063.4; DB 6; Length 2457;  
Best Local Similarity 99.4%; Pred. No. 2.7e-289;  
Matches 1067; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
QY 56 CCCAGCGCGCTCTGAGGAAACAGTACTGCTACTTTGGGAATGGGTGAGCTACCGGTG 115  
Db 627 CCCCTGCTCTGAGGAAACAGTACTGCTACTTTGGGAATGGGTGAGCTACCGGTG 686  
QY 116 GCACGACAGCTTACACGACAGAACCCAGTCCGAGCACTGGGCTGGGCAACATA 175  
Db 687 GCACGACAGCTTACACGACAGAACCCAGTCCGAGCACTGGGCTGGGCAACATA 746  
QY 176 TAGGCAAGTTTACACGACAGAACCCAGTCCGAGCACTGGGCTGGGCAACATA 235  
Db 747 TAGGCAAGTTTACACGACAGAACCCAGTCCGAGCACTGGGCTGGGCAACATA 806  
QY 236 ATTACTCGGGAATCCTGATGGGATGCAAGCCCTGGTGGCAGCTGTGAAGAACCGCA 295  
Db 807 ATTACTCGGGAATCCTGATGGGATGCAAGCCCTGGTGGCAGCTGTGAAGAACCGCA 866  
QY 296 GGCTGACGTGGAGTACTGTGATGTCCTCTGCTCCACCTGCGGCTGAGACAGTACA 355

Db 867 GGCTGACGTGGAGTACTGTGATGTCCTCTGCTCCACCTGCGGCTGAGACAGTACA 926  
QY 356 GCCAGCTTCAGTTTCGATCAAGAGAGGCTCTTCGCGGACATCGCTCCACCCCTGGC 415  
Db 927 GCCAGCTTCAGTTTCGATCAAGAGAGGCTCTTCGCGGACATCGCTCCACCCCTGGC 986  
QY 416 AGGCTGCCATCTTTGCCAAGCACAGGAGGTCCGCGGAGAGCGGTTCCTGCGGGGCA 475  
Db 987 AGGCTGCCATCTTTGCCAAGCACAGGAGGTCCGCGGAGAGCGGTTCCTGCGGGGCA 1046  
QY 476 TACTCATCAGCTCTGCTGCTGATTCCTCTGCGGCGCCACTGCTTCAGAGAGAGTTTCGCG 535  
Db 1047 TACTCATCAGCTCTGCTGCTGATTCCTCTGCGGCGCCACTGCTTCAGAGAGAGTTTCGCG 1106  
QY 536 CCCACCACTGACGCTGATCTTGGCGAAGACATACCGGCTGCTCCCTGGCGAGGAGGAGC 595  
Db 1107 CCCACCACTGACGCTGATCTTGGCGAAGACATACCGGCTGCTCCCTGGCGAGGAGGAGC 1166  
QY 596 AGAAATTTGAAGTCCGAAATAATACATTTGTCATAAGGAATTCGATGATGACATTTACGACA 655  
Db 1167 AGAAATTTGAAGTCCGAAATAATACATTTGTCATAAGGAATTCGATGATGACATTTACGACA 1226  
QY 656 ATGACATTCGCTGCTGACGCTGAATTCGATTCGATTCCTCCGCTGTGCCCGAGAGACGCG 715  
Db 1227 ATGACATTCGCTGCTGACGCTGAATTCGATTCGATTCCTCCGCTGTGCCCGAGGAGCAGCG 1286  
QY 716 TGGTCCGACATGCTGCTTCCCGCGGACCTGACGCTCCGAGCTTGGACGAGTGTG 775  
Db 1287 TGGTCCGACATGCTGCTTCCCGCGGACCTGACGCTCCGAGCTTGGACGAGTGTG 1346  
QY 776 AGCTCTCGGCTACGGCAAGCATGAGGCTTTGCTCTCTTTCTATTTCGAGCGGCTGAAGG 835  
Db 1347 AGCTCTCGGCTACGGCAAGCATGAGGCTTTGCTCTCTTTCTATTTCGAGCGGCTGAAGG 1406  
QY 836 AGGCTCATGCTGACGCTGATCCATCCAGCTGACATCAACAATTTACTTAACAGAA 895  
Db 1407 AGGCTCATGCTGACGCTGATCCATCCAGCTGACATCAACAATTTACTTAACAGAA 1466  
QY 896 CAGTCACCGACAACATGCTGTGCTGAGACACTCGGAGCGGGGCGGCGGCAAACT 955  
Db 1467 CAGTCACCGACAACATGCTGTGCTGAGACACTCGGAGCGGGGCGGCGGCAAACT 1526  
QY 956 TGACAGACGCTGCGCAGGCGGATTCGGAGAGCGCCCTGCTGTGCTGAACGATGCGCGCA 1015  
Db 1527 TGACAGACGCTGCGCAGGCGGATTCGGAGAGCGCCCTGCTGTGCTGAACGATGCGCGCA 1586  
QY 1016 TGACTTTGGTGGGATCATCAGCTGGGCTGTGGGCTGTGGACAGAGATGTCGCGGTG 1075  
Db 1587 TGACTTTGGTGGGATCATCAGCTGGGCTGTGGGCTGTGGACAGAGATGTCGCGGTG 1646  
QY 1076 TGTACAAAAGTTTACCAACTACCTAGACTGATTCGTCGACAACTGCGACG 1128  
Db 1647 TGTACAAAAGTTTACCAACTACCTAGACTGATTCGTCGACAACTGCGACG 1699

RESULT 11  
US-08-286-740-1  
; Sequence 1, Application US/08286740  
; Patent No. 5561053  
; GENERAL INFORMATION:  
; APPLICANT: Crowley, Craig W.  
; TITLE OF INVENTION: METHOD FOR SELECTING HIGH-EXPRESSING  
; TITLE OF INVENTION: HOST CELLS  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Genentech, Inc.  
; STREET: 460 Point San Bruno Blvd  
; CITY: South San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94080  
; COMPUTER READABLE FORM:



MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: patin (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/286,740  
FILING DATE: 05-AUG-1994  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Lee, Wendy M.  
REGISTRATION NUMBER:  
REFERENCE/DOCKET NUMBER: 798  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415/225-1994  
TELEFAX: 415/952-9881  
TELEX: 910/371-7168  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 7360 bases  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
US-08-286-740-1

Query Match 94.3%; Score 1063.4; DB 1; Length 7360;  
Best Local Similarity 99.4%; Pred. No. 4.3e-289;  
Matches 1067; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 56 CCACGCGCGCTCTCAGGAAACAGTACTTCTTTGGGAATGGGTGAGCTACCGTG 115  
DB 2235 CCCCCTGCTCTCTAGGGAACAGTACTTCTTTGGGAATGGGTGAGCTACCGTG 2294  
QY 116 GCACGACAGCTCACCAGTGGGTGCTCTGCTCCCGTGGGAATTCATGATCTGA 175  
DB 2295 GCACGACAGCTCACCAGTGGGTGCTCTGCTCCCGTGGGAATTCATGATCTGA 2354  
QY 176 TAGGCAAGTTTACACACACAGAACCCAGTCCCGAGCACTGGGCTTGGCAACATA 235  
DB 2355 TAGGCAAGTTTACACACACAGAACCCAGTCCCGAGCACTGGGCTTGGCAACATA 2414  
QY 236 ATTACTGCGGGAATCTGATGGGATGCCAGCCCTGTGTCACGTCTGAAGAACGCA 295  
DB 2415 ATTACTGCGGGAATCTGATGGGATGCCAGCCCTGTGTCACGTCTGAAGAACGCA 2474  
QY 296 GGCTGACGTGGAGTACTGTGATGGCCCTCTGCTCCACCTGCGGCTTGAGACAGTACA 355  
DB 2475 GGCTGACGTGGAGTACTGTGATGGCCCTCTGCTCCACCTGCGGCTTGAGACAGTACA 2534  
QY 356 GCCAGCCTCAGTTTGGCATCAAGAGGGCTCTTGGCGGATCGCCTCCACCCCTGGC 415  
DB 2535 GCCAGCCTCAGTTTGGCATCAAGAGGGCTCTTGGCGGATCGCCTCCACCCCTGGC 2594  
QY 416 AGGCTGCCATCTTCCCAAGCACAGAGTGGCCGAGAGCGGTTCCTGTGCGGGGCA 475  
DB 2595 AGGCTGCCATCTTCCCAAGCACAGAGTGGCCGAGAGCGGTTCCTGTGCGGGGCA 2654  
QY 476 TACTCATCAGCTCTGTGTTGATTTCTCTGCGGCCACTGTCTTCAGGAGAGTTTCGCG 535  
DB 2655 TACTCATCAGCTCTGTGTTGATTTCTCTGCGGCCACTGTCTTCAGGAGAGTTTCGCG 2714  
QY 536 CCCACCACTGACGTGTATCTTGGCAACATACCGGGTGGTCCCTGGCGAGGAGGAGC 595  
DB 2715 CCCACCACTGACGTGTATCTTGGCAACATACCGGGTGGTCCCTGGCGAGGAGGAGC 2774  
QY 596 AGAAATTTGAAGTCGAAAAATACATTGTCCATAGGAATTCATGATCAGACTTACGACA 655  
DB 2775 AGAAATTTGAAGTCGAAAAATACATTGTCCATAGGAATTCATGATCAGACTTACGACA 2834  
QY 656 ATGACATTCGCTGCTGACGTGAAATCGGATTCGTCCTCGCTGTGCCAGGAGACGACG 715

DB 2835 ATGACATTCGCTGCTGACGTGAAATCGGATTCGTCCTCGCTGTGCCAGGAGACGCG 2894  
QY 716 TGGTCCGCACTGTGTGCTTCCCGCGGAGCACTGCGAGCTGCGGACTTGGACGAGTGTG 775  
DB 2895 TGGTCCGCACTGTGTGCTTCCCGCGGAGCACTGCGAGCTGCGGACTTGGACGAGTGTG 2954  
QY 776 AGCTCTCCGCTACGGCAAGCATGAGGCTTGTCTCTCTTTCTATTTCGAGCGGCTGAAGG 835  
DB 2955 AGCTCTCCGCTACGGCAAGCATGAGGCTTGTCTCTCTTTCTATTTCGAGCGGCTGAAGG 3014  
QY 836 AGGCTCATGTGACACTGTATCCATCCAGCGCTGCAATCACAAATTTACTTAACAGAA 895  
DB 3015 AGGCTCATGTGACACTGTATCCATCCAGCGCTGCAATCACAAATTTACTTAACAGAA 3074  
QY 896 CAGTCCGCAACATCTGTGCTGAGACACTCGGAGCGGCGGCCAGGCAAACT 955  
DB 3075 CAGTCCGCAACATCTGTGCTGAGACACTCGGAGCGGCGGCCAGGCAAACT 3134  
QY 956 TGCAACGACGCTGCGAGGCGAATTCGGGAGGCGCCCTGGTGTGTCTGAACGATGCCGCA 1015  
DB 3135 TGCAACGACGCTGCGAGGCGAATTCGGGAGGCGCCCTGGTGTGTCTGAACGATGCCGCA 3194  
QY 1016 TGACTTTGGTGGGATCATCATGCTGGGCGCTTGGGCTGTGACAGAGATGTCGCGGTG 1075  
DB 3195 TGACTTTGGTGGGATCATCATGCTGGGCGCTTGGGCTGTGACAGAGATGTCGCGGTG 3254  
QY 1076 TGACACAAAGTTTACCAACTTACCTAGACTGGATTTCGTGACAACTCGACCG 1128  
DB 3255 TGACACAAAGTTTACCAACTTACCTAGACTGGATTTCGTGACAACTCGACCG 3307

RESULT 12  
PCT-US95-09576-1  
; Sequence 1, Application PC/TUS9509576  
; GENERAL INFORMATION:  
; APPLICANT: GENENTECH, INC.  
; TITLE OF INVENTION: METHOD FOR SELECTING HIGH-EXPRESSING  
; TITLE OF INVENTION: HOST CELLS  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Genentech, Inc.  
; STREET: 460 Point San Bruno Blvd  
; CITY: South San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94080  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: patin (Genentech)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US95/09576  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/286740  
; FILING DATE: 05-AUG-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Lee, Wendy M.  
; REGISTRATION NUMBER: 00,000  
; REFERENCE/DOCKET NUMBER: 798PCT  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415/225-1994  
; TELEFAX: 415/952-9881  
; TELEX: 910/371-7168  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 7360 bases  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
PCT-US95-09576-1







596 AGAAATTTGAAGTCGAAAAATACATTTGTCATAGAAATTCGATGATGACACTTACGACA 655  
779 AGAAATTTGAAGTCGAAAAATACATTTGTCATAGAAATTCGATGATGACACTTACGACA 838  
656 ATGACATTTGCGCTGCTGACGATGAATCGGATTCGTCGCTGCTGCCAGGACGAGG 715  
839 ATGACATTTGCGCTGCTGACGATGAATCGGATTCGTCGCTGCTGCCAGGACGAGG 898  
716 TGGTCCGCACTGTGTGCTTCCCGCGGACCTGCGAGCTGCGGACCTGCGAGGAGTG 775  
899 TGGTCCGCACTGTGTGCTTCCCGCGGACCTGCGAGCTGCGGACCTGCGAGGAGTG 958  
776 AGCTCTCCGCTACGCGAAGCATGAGGCTTGTCTCTCTTCTATTCGAGGCGCTGAAGG 835  
959 AGCTCTCCGCTACGCGAAGCATGAGGCTTGTCTCTCTTCTATTCGAGGCGCTGAAGG 1018  
836 AGGCTCATGTGACATGTATCCCATCCAGCGCTGACATCAACATTTACTTACAGAA 895  
1019 AGGCTCATGTGACATGTATCCCATCCAGCGCTGACATCAACATTTACTTACAGAA 1078  
896 CAGTCACCGACAACATGTGTGTGCTGAGACACTCGAGCGCGCGGCCCGCAAACT 955  
1079 CAGTCACCGACAACATGTGTGTGCTGAGACACTCGAGCGCGCGGCCCGCAAACT 1138  
956 TGCACGAGCTGCGCAGGCGATTCGAGGCGCCCTGCTGTGCTGAACGATGCCGCA 1015  
1139 TGCACGAGCTGCGCAGGCGATTCGAGGCGCCCTGCTGTGCTGAACGATGCCGCA 1198  
1016 TGACTTTGGTGGGCATCATCAGCTGGGCGCTGGCTGTGGACAGAGATGTCCGGGTG 1075  
1199 TGACTTTGGTGGGCATCATCAGCTGGGCGCTGGCTGTGGACAGAGATGTCCGGGTG 1258  
1076 TGTACAAAAGTTTACCACTACCTAGACTGATTCGTAACATGCGACCG 1128  
1259 TGTACAAAAGTTTACCACTACCTAGACTGATTCGTAACATGCGACCG 1311

## RESULT 15

US-08-811-949-38  
Sequence 38, Application US/08811949  
Patent No. 5840533  
GENERAL INFORMATION:  
APPLICANT: NIWA, MINEO  
APPLICANT: SAITO, YOSHIMASA  
APPLICANT: SASAKI, HITOSHI  
APPLICANT: HAYASHI, MASAKO  
APPLICANT: NOTANI, JOUJI  
APPLICANT: KOBAYASHI, MASAKAZU  
TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR  
NUMBER OF SEQUENCES: 67  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,  
ADDRESSEE: P.C.  
STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400  
CITY: ARLINGTON  
STATE: VA  
COUNTRY: USA  
ZIP: 22202  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/811,949  
FILING DATE: 05-MAR-1997  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: OBLON, NORMAN F.  
REGISTRATION NUMBER: 24,618  
REFERENCE/DOCKET NUMBER: 18-966-0  
TELECOMMUNICATION INFORMATION:

TELEPHONE: 703-413-3000  
TELEFAX: 703-413-2220  
INFORMATION FOR SEQ ID NO: 38:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1974 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: circular  
MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 3..1583  
US-08-811-949-38

Query Match 94.1%; Score 1061.8; DB 2; Length 1974;  
Best Local Similarity 99.3%; Pred. No. 7e-289;  
Matches 1066; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 56 CCCAGGCGGCTCTGAGGGAACAGTGACTCTACTTTGGGAATGGTTCAGCCTACCGTG 115  
Db 511 CCCCTGCTGCTCTGAGGGAACAGTGACTCTACTTTGGGAATGGTTCAGCCTACCGTG 570  
QY 116 GCAGCAGAGCTCACCGAGTCGGGTGCTCTGCTCCCTGGAATTCATGATCCTGA 175  
Db 571 GCAGCAGAGCTCACCGAGTCGGGTGCTCTGCTCCCTGGAATTCATGATCCTGA 630  
QY 176 TAGGCAAGGTTTACACAGCAGAACCCAGTCGCCAGGCACTGGGCTGGGCAACATA 235  
Db 631 TAGGCAAGGTTTACACAGCAGAACCCAGTCGCCAGGCACTGGGCTGGGCAACATA 690  
QY 236 ATTAAGTCCGGAATCTGATGGGATGCCAAGCCCTGGTGCAGCTGTGAAGAACCGCA 295  
Db 691 ATTAAGTCCGGAATCTGATGGGATGCCAAGCCCTGGTGCAGCTGTGAAGAACCGCA 750  
QY 296 GGCTGAGCTGGAGTACTGTGATGCTCCCTGCTCCACCTGGGCTGGAGACAGTACA 355  
Db 751 GGCTGAGCTGGAGTACTGTGATGCTCCCTGCTCCACCTGGGCTGGAGACAGTACA 810  
QY 356 GCCAGCTCAGTTTCGATCAAGAGGGCTCTTTCGCCGACATGCCCTCCCAACCCCTGGC 415  
Db 811 GCCAGCTCAGTTTCGATCAAGAGGGCTCTTTCGCCGACATGCCCTCCCAACCCCTGGC 870  
QY 416 AGGTGCGCATCTTTGCCAAGCAGAGAGTCCGCCGAGAGCGGTTCCTGTGCGGGGCA 475  
Db 871 AGGTGCGCATCTTTGCCAAGCAGAGAGTCCGCCGAGAGCGGTTCCTGTGCGGGGCA 930  
QY 476 TACTCATCAGCTCCTGTGGATTCCTCTGCGGCGCCACTGCTTCAGAGAGGTTTCGC 535  
Db 931 TACTCATCAGCTCCTGTGGATTCCTCTGCGGCGCCACTGCTTCAGAGAGGTTTCGC 990  
QY 536 CCCACCACTGACGCTGATCTTGGGCAAGAACATACCGGGTGGTCCCTGGCGAGGAGGAGC 595  
Db 991 CCCACCACTGACGCTGATCTTGGGCAAGAACATACCGGGTGGTCCCTGGCGAGGAGGAGC 1050  
QY 596 AGAAATTTGAAGTCGAAAAATACATTTGTCATAGAAATTCGATGATGACACTTACGACA 655  
Db 1051 AGAAATTTGAAGTCGAAAAATACATTTGTCATAGAAATTCGATGATGACACTTACGACA 1110  
QY 656 ATGACATTTGCGCTGCTGACGATGAATCGGATTCGTCGCTGCTGCCAGGAGGAGCG 715  
Db 1111 ATGACATTTGCGCTGCTGACGATGAATCGGATTCGTCGCTGCTGCCAGGAGGAGCG 1170  
QY 716 TGGTCCGCACTGTGTGCTTCCCGCGGACCTGCGAGCTGCGGACCTGGAAGGAGTG 775  
Db 1171 TGGTCCGCACTGTGTGCTTCCCGCGGACCTGCGAGCTGCGGACCTGGAAGGAGTG 1230  
QY 776 AGCTCTCCGCTACGCGAAGCATGAGGCTTGTCTCTCTTCTATTCGAGGCGCTGAAGG 835  
Db 1231 AGCTCTCCGCTACGCGAAGCATGAGGCTTGTCTCTCTTCTATTCGAGGCGCTGAAGG 1290  
QY 836 AGGCTCATGTGACATGTATCCCATCCAGCGCTGACATCAACATTTACTTACAGAA 895  
Db 1291 AGGCTCATGTGACATGTATCCCATCCAGCGCTGACATCAACATTTACTTACAGAA 1350

Qy	896	CAGTCAACCGACAAACATGCTGTGTGCTGGAGACACTCGGAGCGCGGGCCCCCAGGCAAACT	955
Db	1351		
Qy	956	TGCACGACGCTGCGCAGGGCGATTGCGGAGGCCCCCTGGTGTCTGAAACGATGGCCGCA	1015
Db	1411		
Qy	1016	TGACTTTGGTGGGCATCATCAGCTGGGGCCTGGGCTGTGGACAGGAGATGTCCCGGTG	1075
Db	1471		
Qy	1076	TGTACACAAAGTTTACCACTACCTAGACTGGATTGCTGACAAACATGCGACCG	1128
Db	1531		
		TGTACACAAAGTTTACCACTACCTAGACTGGATTGCTGACAAACATGCGACCG	1583

Search completed: August 19, 2003, 08:35:06  
Job time : 90.2804 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 19, 2003, 04:46:44 ; Search time 318.307 Seconds  
(without alignments)  
7930.701 Million cell updates/sec

Title: US-09-987-455-2

Perfect score: 1128

Sequence: 1 atgaaaagacagtatcgc.....ttcgtgacaacatgcgacgcg 1128

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 1504479 seqs, 1118970152 residues

Total number of hits satisfying chosen parameters: 3008958

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA:\*

- 1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq:\*
- 3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:\*
- 4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:\*
- 5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:\*
- 6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq:\*
- 7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:\*
- 8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*
- 9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq:\*
- 10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq:\*
- 11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq:\*
- 12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:\*
- 13: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq:\*
- 14: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq:\*
- 15: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:\*
- 16: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:\*
- 17: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1128	100.0	1128	11	US-09-987-455-2
2	1128	100.0	1128	11	US-09-987-455-5
3	1063.4	94.3	1689	9	US-09-969-271-6
4	1063.4	94.3	2509	14	US-10-193-656-7
5	1063.4	94.3	2519	9	US-09-969-271-5
6	1062	94.1	1065	11	US-09-987-455-4
7	1062	94.1	1065	11	US-09-987-455-7
8	1061.8	94.1	2641	10	US-09-974-298-144
9	227.2	20.1	329	12	US-10-007-926A-433
10	191.6	17.0	1212	10	US-09-880-503-15
11	188	16.7	1415	14	US-10-198-846-12748
12	170.8	15.1	1475	9	US-09-735-705-122
13	170.8	15.1	1475	10	US-09-850-716A-122
14	170.8	15.1	1475	10	US-09-897-778-122
15	170.8	15.1	1475	11	US-09-466-396A-122
16	170.8	15.1	1475	12	US-10-117-982-122

Sequence 159, App  
Sequence 46, Appl  
Sequence 458, App  
Sequence 1, Appli  
Sequence 123, App  
Sequence 123, App  
Sequence 123, App  
Sequence 123, App  
Sequence 12, Appl  
Sequence 160, App  
Sequence 183, App  
Sequence 1682, Ap  
Sequence 1989, Ap  
Sequence 14, Appl  
Sequence 16, Appl  
Sequence 10405, A  
Sequence 743, App  
Sequence 448, App  
Sequence 1, Appli  
Sequence 552, App  
Sequence 1612, Ap  
Sequence 18438, A  
Sequence 11188, A  
Sequence 237, App  
Sequence 217, App  
Sequence 12715, A  
Sequence 35, Appl

17 170.8 15.1 1475 12 US-10-101-510-159  
18 170.8 15.1 2336 12 US-09-971-392-46  
19 170.8 15.1 2336 12 US-10-101-510-458  
20 167.6 14.9 1296 13 US-10-076-421-1  
21 167.6 14.9 2294 9 US-09-735-705-123  
22 167.6 14.9 2294 10 US-09-850-716A-123  
23 167.6 14.9 2294 10 US-09-897-778-123  
24 167.6 14.9 2294 11 US-09-466-396A-123  
25 167.6 14.9 2294 12 US-10-117-982-123  
26 167.6 14.9 2344 12 US-10-101-510-383  
27 166 14.7 1236 10 US-09-880-503-12  
28 166 14.7 2304 12 US-10-301-822-160  
29 166 14.7 2304 14 US-10-171-311-183  
30 162 14.5 474 9 US-09-864-761-1682  
31 162 14.4 2486 14 US-10-106-698-1989  
32 159.8 14.2 831 10 US-09-880-503-14  
33 159.8 14.2 972 10 US-09-880-503-16  
34 139 12.3 461 11 US-09-918-995-10405  
35 123 10.9 1959 12 US-10-101-510-743  
36 123 10.9 2035 12 US-10-101-510-448  
37 123 10.9 2048 11 US-09-858-909-1  
38 120.4 10.7 2036 10 US-09-954-456-552  
39 120.4 10.7 2036 10 US-09-880-107-1612  
40 111.4 9.9 116 9 US-09-864-761-18438  
41 105 9.3 457 11 US-09-918-995-11188  
C 42 89.4 7.9 372 10 US-09-520-300A-237  
C 43 89.4 7.9 372 13 US-10-033-528-237  
C 44 88 7.8 442 11 US-09-918-995-12715  
45 83.8 7.4 1230 9 US-09-879-792-35

#### ALIGNMENTS

#### RESULT 1

US-09-987-455-2  
; Sequence 2, Application US/09987455  
; Publication No. US200300049729A1  
; GENERAL INFORMATION:  
; APPLICANT: Aranya Manosroi  
; APPLICANT: Jiradej Manosroi  
; APPLICANT: Chatchai Tayasawatana  
; APPLICANT: Friedrich Goetz  
; APPLICANT: Rolf-Guenther Werner  
; TITLE OF INVENTION: Methods for Large Scale Production of Recombinant  
; TITLE OF INVENTION: DNA-Derived tPA or K2S Molecules  
; FILE REFERENCE: 0652.2190001  
; CURRENT APPLICATION NUMBER: US/09/987,455  
; CURRENT FILING DATE: 2001-11-14  
; PRIOR APPLICATION NUMBER: 60/268,574  
; PRIOR FILING DATE: 2001-02-15  
; PRIOR APPLICATION NUMBER: GB 0027779.8  
; PRIOR FILING DATE: 2000-11-14  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 2

LENGTH: 1128

TYPE: DNA

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Description of Artificial Sequence: coding

OTHER INFORMATION: sequence for OmpA-K2S fusion protein

US-09-987-455-2

Query Match 100.0%; Score 1128; DB 11; Length 1128;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1128; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGAAAAGACAGCTATCGGATTCGAGTGCAGTGGTTCGTTACGTCGCCAG 60

Db 1 ATGAAAAGACAGCTATCGGATTCGAGTGCAGTGGTTCGTTACGTCGCCAG 60

Qy 61 CGGCGCTCTGAGGGAACAGTACTGCTTTTGGGAATGGGTACGCTACCGTGGCAG 120



Qy	601	TTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAATGAC	660
Db	601	TTTGAAGTCGAAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGACAATGAC	660
Qy	661	ATTGCGCTGCTGCAGCTGMAATCGGATTGCTCCGCTGTGCCAGGAGACGCGTGGTC	720
Db	661	ATTGCGCTGCTGCAGCTGMAATCGGATTGCTCCGCTGTGCCAGGAGACGCGTGGTC	720
Qy	721	CGCACTGTGTGCCCTTCCCCCGGGCGAAGCTGCAGCTGCGGGACTGGACGGAGTGTGAGCTC	780
Db	721	CGCACTGTGTGCCCTTCCCCCGGGCGAAGCTGCAGCTGCGGGACTGGACGGAGTGTGAGCTC	780
Qy	781	TCGGGCTACGGCAAGCATGAGGCTTGCTCTCTTTATTTCGGAGCGGCTGAAGAGGCT	840
Db	781	TCGGGCTACGGCAAGCATGAGGCTTGCTCTCTTTATTTCGGAGCGGCTGAAGAGGCT	840
Qy	841	CATGTGAGCTGTACCATTCAGCGCTGCACATCACAACTTTACTTTACAGAACAGTC	900
Db	841	CATGTGAGCTGTACCATTCAGCGCTGCACATCACAACTTTACTTTACAGAACAGTC	900
Qy	901	ACGCACAACATGCTGTGCTGGAGACACTCGGAGCGGCGGCCCCAGGCGAAACTTCAC	960
Db	901	ACGCACAACATGCTGTGCTGGAGACACTCGGAGCGGCGGCCCCAGGCGAAACTTCAC	960
Qy	961	GAGCCTTCGCAGGGCGAATTCGGGAGGCCCTCTGTGTGTCTGAACGATGCGCGATGACT	1020
Db	961	GAGCCTTCGCAGGGCGAATTCGGGAGGCCCTCTGTGTGTCTGAACGATGCGCGATGACT	1020
Qy	1021	TTGTGTGGGCATCATCAGCTGGGGCTTGGGCTGTGGACAGAGGATGTCCCGGGTGTGTAC	1080
Db	1021	TTGTGTGGGCATCATCAGCTGGGGCTTGGGCTGTGGACAGAGGATGTCCCGGGTGTGTAC	1080
Qy	1081	ACAAAGTTTACCAACTACCTAGACTGGATTTCGTGACCAACTGCGACCG	1128
Db	1081	ACAAAGTTTACCAACTACCTAGACTGGATTTCGTGACCAACTGCGACCG	1128

### RESULT 3

```

US-09-969-271-6
; Sequence 6, Application US/09969271
; Patent No. US20020098179A1
; GENERAL INFORMATION:
; APPLICANT: Pfizer Inc. (All designated States except GB and EP (GB));
; APPLICANT: Pfizer Limited (GB and EP (GB) only)
; TITLE OF INVENTION: Pharmaceutical Combinations
; FILE REFERENCE: PCS1095IAPME
; CURRENT APPLICATION NUMBER: US/09/969,271
; CURRENT FILING DATE: 2001-10-01
; PRIOR APPLICATION NUMBER: GB 0025473.0
; PRIOR FILING DATE: 2000-10-17
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 1689
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-969-271-6

```

## RESULT 4

RESULT 4  
 US-10-193-656-7  
 : Sequence 7, Application US/10193656  
 : Publication No. US20030096733A1  
 : GENERAL INFORMATION:  
 : APPLICANT: NY, Tor  
 : APPLICANT: HOLMDEHL, Rikard  
 : APPLICANT: Li, Jinan  
 : TITLE OF INVENTION: NOVEL DRUG TARGETS FOR ARTHRITIS  
 : FILE REFERENCE: 3810/1J577-US3  
 : CURRENT APPLICATION NUMBER: US/10/193,656





Db 1050 AGGTCGCACTTTTCCCAAGCACAGGAGTGCCTCCGAGAGCGTTCCTGTGCGGGGCA 1109  
Qy 476 TACTCATCAGCTCCTGCTGGATTCTCTGCGGCCACTGCTTCCAGGAGGTTTCCGC 535  
Db 1110 TACTCATCAGCTCCTGCTGGATTCTCTGCGGCCACTGCTTCCAGGAGGTTTCCGC 1169  
Qy 536 CCCACCACTGACGCTGATCTTTGGGCAAGAACATACCGGGTGTCTCCGCGAGGAGGAGC 595  
Db 1170 CCCACCACTGACGCTGATCTTTGGGCAAGAACATACCGGGTGTCTCCGCGAGGAGGAGC 1229  
Qy 596 AGAATTTGAAGTCGAAATATACATTTGTCATAGGAATTCGATGATGACACTTACGACA 655  
Db 1230 AGAATTTGAAGTCGAAATATACATTTGTCATAGGAATTCGATGATGACACTTACGACA 1289  
Qy 656 ATGACATTTGCGCTGCTGACGCTGAAATCGGATTCGTCCTCCGCTGTGCCAGGAGGAGCAGC 715  
Db 1290 ATGACATTTGCGCTGCTGACGCTGAAATCGGATTCGTCCTCCGCTGTGCCAGGAGGAGCAGC 1349  
Qy 716 TGGTCCGCACTGTGTGCTTCCCGCGGACCTGACGCTGCGGAGCTGGAACGAGTGTG 775  
Db 1350 TGGTCCGCACTGTGTGCTTCCCGCGGACCTGACGCTGCGGAGCTGGAACGAGTGTG 1409  
Qy 776 AGCTCTCCGCTACGCGAAGCATGAGGCTTGTCTCTTTCTATTTCGAGCGGCTGAGG 835  
Db 1410 AGCTCTCCGCTACGCGAAGCATGAGGCTTGTCTCTTTCTATTTCGAGCGGCTGAGG 1469  
Qy 836 AGGCTCATGTGACGCTGACCCATCCAGCGCTGCACATCAACATTTACTTAACAGAA 895  
Db 1470 AGGCTCATGTGACGCTGACCCATCCAGCGCTGCACATCAACATTTACTTAACAGAA 1529  
Qy 896 CAGTCACCGCAACATGTGTGTGTGAGACACTCGGAGCGGGGCCCCAGGCAAACT 955  
Db 1530 CAGTCACCGCAACATGTGTGTGTGAGACACTCGGAGCGGGGCCCCAGGCAAACT 1589  
Qy 956 TGCAGACGCTGCCAGGCGGATTCGGAGGCGCCCTGCTGTGTCTGAAAGATGTCGCGGTG 1015  
Db 1590 TGCAGACGCTGCCAGGCGGATTCGGAGGCGCCCTGCTGTGTCTGAAAGATGTCGCGGTG 1649  
Qy 1016 TGACTTTGCTGGGATCATAGCTGGGCTGGGCTGGGAGAGGATGTCGCGGTG 1075  
Db 1650 TGACTTTGCTGGGATCATAGCTGGGCTGGGCTGGGAGGATGTCGCGGTG 1709  
Qy 1076 TGTACACAAAGTTACCAACTACCTAGACTGGATTTCGTGACAAATGCGACCG 1128  
Db 1710 TGTACACAAAGTTACCAACTACCTAGACTGGATTTCGTGACAAATGCGACCG 1762

RESULT 6

US-09-987-455-4  
; Sequence 4, Application US/09987455  
; Publication No. US20030049729A1  
; GENERAL INFORMATION:  
; APPLICANT: Aranya Manosroi  
; APPLICANT: Jiradej Manosroi  
; APPLICANT: Chatchai Tayapiwatana  
; APPLICANT: Friedrich Goetz  
; APPLICANT: Rolf-Guenther Werner  
; TITLE OF INVENTION: Methods for Large Scale Production of Recombinant  
; FILE OF INVENTION: DNA-Derived tPA or K2S Molecules  
; FILE REFERENCE: 0652.2190001  
; CURRENT APPLICATION NUMBER: US/09/987,455  
; CURRENT FILING DATE: 2001-11-14  
; PRIOR FILING DATE: 60/268,574  
; PRIOR FILING DATE: 2001-02-15  
; PRIOR APPLICATION NUMBER: GB 002779.8  
; PRIOR FILING DATE: 2000-11-14  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 4  
; LENGTH: 1065  
; TYPE: DNA  
; ORGANISM: Artificial Sequence

; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: coding  
; OTHER INFORMATION: sequence for K2S protein  
US-09-987-455-4  
Query Match 94.18; Score 1062; DB 11; Length 1065;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1062; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 67 TCTGAGGAAACAGTGTCTTCTTTGGGAAATGGGTGAGCTTACCGTGGCAGCACAGC 126  
Db 1 TCTGAGGAAACAGTGTCTTCTTTGGGAAATGGGTGAGCTTACCGTGGCAGCACAGC 60  
Qy 127 CTCAACGAGTGGGTGCTCTCTGCTCCCGTGGAAATTCATGATCTCTGATAGGCAAGTT 186  
Db 61 CTCAACGAGTGGGTGCTCTCTGCTCCCGTGGAAATTCATGATCTCTGATAGGCAAGTT 120  
Qy 187 TACACACACAGAACCCAGTGCAGGCACTGGGCTGGGCAACATTAATTAATCTGCGG 246  
Db 121 TACACACACAGAACCCAGTGCAGGCACTGGGCTGGGCAACATTAATTAATCTGCGG 180  
Qy 247 AATCTGATGGGATGCAAGCCCTGCTGCGCACGCTGCTGAAGAACCCGAGGCTGACGTG 306  
Db 181 AATCTGATGGGATGCAAGCCCTGCTGCGCACGCTGCTGAAGAACCCGAGGCTGACGTG 240  
Qy 307 GAGTACTGTGATGTCCTCTCTGCTCCACTGCGGCTGAGACAGTACAGCCCTCAG 366  
Db 241 GAGTACTGTGATGTCCTCTCTGCTCCACTGCGGCTGAGACAGTACAGCCCTCAG 300  
Qy 367 TTTGCGATCAAGAGGAGGCTTTCGCGACATCGCTCCACCCCTGGCAGGCTGCCATC 426  
Db 301 TTTGCGATCAAGAGGAGGCTTTCGCGACATCGCTCCACCCCTGGCAGGCTGCCATC 360  
Qy 427 TTTGCCAAGCACAGGAGTGCCTGCGGAGAGCGGTTCCTGTCGGGGGCGATCTCATCAGC 486  
Db 361 TTTGCCAAGCACAGGAGTGCCTGCGGAGAGCGGTTCCTGTCGGGGGCGATCTCATCAGC 420  
Qy 487 TCCTGCTGGATTCTCTCTGCGGCCACTGCTTCAGGAGAGGTTTCCGCCCCACCATCTG 546  
Db 421 TCCTGCTGGATTCTCTCTGCGGCCACTGCTTCAGGAGAGGTTTCCGCCCCACCATCTG 480  
Qy 547 ACGTGATCTTTGGGCAAGAACATACCGGGTGTCTCCGCGAGGAGGAGAGAAATTTGAA 606  
Db 481 ACGTGATCTTTGGGCAAGAACATACCGGGTGTCTCCGCGAGGAGGAGAGAAATTTGAA 540  
Qy 607 GTCCAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGAAATGACATTTGCG 666  
Db 541 GTCCAAAATACATTGTCCATAAGGAATTCGATGATGACACTTACGAAATGACATTTGCG 600  
Qy 667 CTGCTGAGCTGAAATCGGATTCGTCCTCCGCTGTGCCAGGAGGAGCGTGGTCCGCACT 726  
Db 601 CTGCTGAGCTGAAATCGGATTCGTCCTCCGCTGTGCCAGGAGGAGCGTGGTCCGCACT 660  
Qy 727 GTGTGCTTCCCGCGGAGCTGACGCTGCGGAGTGCAGCGGAGTGTGAGCTTCCGCG 786  
Db 661 GTGTGCTTCCCGCGGAGCTGACGCTGCGGAGTGCAGCGGAGTGTGAGCTTCCGCG 720  
Qy 787 TACGCAAGCATGAGGCTTGTCTCTTCTATTTCGAGCGGCTGAAAGGAGGCTCATGTC 846  
Db 721 TACGCAAGCATGAGGCTTGTCTCTTCTATTTCGAGCGGCTGAAAGGAGGCTCATGTC 780  
Qy 847 AGACTGTACCATCCAGCCGCTGCACATCAACATTTACTTAACAGAACAGTCCAGCAG 906  
Db 781 AGACTGTACCATCCAGCCGCTGCACATCAACATTTACTTAACAGAACAGTCCAGCAG 840  
Qy 907 AACATGCTGTGTGTGAGACACTCGGAGCGGCGGCCCCAGGCAAACTTGTGACAGGCG 966  
Db 841 AACATGCTGTGTGTGAGACACTCGGAGCGGCGGCCCCAGGCAAACTTGTGACAGGCG 900  
Qy 967 TGCAGGCGGATTCGGGAGGCGGCCCCCTGCTGTAACGATGCGCCGCTGCTTGGTG 1026  
Db 901 TGCAGGCGGATTCGGGAGGCGGCCCCCTGCTGTAACGATGCGCCGCTGCTTGGTG 960

QY 1027 GGCATCATCAGCTGGGGCTTGGCTGTGGACAGAGGATGTCCCGGTGTGTACACAAAG 1086  
Db 961 GGCATCATCAGCTGGGGCTTGGCTGTGGACAGAGGATGTCCCGGTGTGTACACAAAG 1020  
QY 1087 GTTACCAACTACCTAGACTGGATTCTGTGACACATGCGACCG 1128  
Db 1021 GTTACCAACTACCTAGACTGGATTCTGTGACACATGCGACCG 1062

## RESULT 7

US-09-987-455-7  
; Sequence 7, Application US/09987455  
; Publication No. US20030049729A1  
; GENERAL INFORMATION:  
; APPLICANT: Aranya Manosroi  
; APPLICANT: Jiradej Manosroi  
; APPLICANT: Chatchai Tayapiwatana  
; APPLICANT: Friedrich Goetz  
; TITLE OF INVENTION: Methods for Large Scale Production of Recombinant  
; TITLE OF INVENTION: DNA-Derived tpa or K2S Molecules  
; FILE REFERENCE: 0652.2190001  
; CURRENT APPLICATION NUMBER: US/09/987,455  
; PRIOR FILING DATE: 2001-11-14  
; PRIOR APPLICATION NUMBER: 60/268,574  
; PRIOR FILING DATE: 2001-02-15  
; PRIOR APPLICATION NUMBER: GB 0027779.8  
; PRIOR FILING DATE: 2000-11-14  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 7  
; LENGTH: 1065  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: coding  
; OTHER INFORMATION: sequence for K2S protein  
US-09-987-455-7

Query Match 94.1%; Score 1062; DB 11; Length 1065;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1062; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 67 TCTGAGGAAACAGTACTGCTTCTTGGGAATGGTCTAGCTACCGTGGCAGCAGCAGC 126  
Db 1 TCTGAGGAAACAGTACTGCTTCTTGGGAATGGTCTAGCTACCGTGGCAGCAGCAGC 60  
QY 127 CTCACCGAGTGGGTGCTCTCTGCTCCCGTGGAAATTCATGATCCTGATAGGCAAGTT 186  
Db 61 CTCACCGAGTGGGTGCTCTCTGCTCCCGTGGAAATTCATGATCCTGATAGGCAAGTT 120  
QY 187 TACACAGCAGAAACCCAGTGGCCAGCAGTGGCCCTGGGCAACATTAATCTGCGG 246  
Db 121 TACACAGCAGAAACCCAGTGGCCAGCAGTGGCCCTGGGCAACATTAATCTGCGG 180  
QY 247 AATCTGATGGGATGCCAAGCCCTGTGTGTCACAGTGTGGAAGCCGAGCTGACGTGG 306  
Db 181 AATCTGATGGGATGCCAAGCCCTGTGTGTCACAGTGTGGAAGCCGAGCTGACGTGG 240  
QY 307 GAGTACTGTGATGTCCTCTCTGCTCCACCTGCGGCTTGAGACAGTACAGCCCTCAG 366  
Db 241 GAGTACTGTGATGTCCTCTCTGCTCCACCTGCGGCTTGAGACAGTACAGCCCTCAG 300  
QY 367 TTTCCATCAAGAGGGCTCTTCCGACATCGCTCCACCCCTGCGAGCTGCCATC 426  
Db 301 TTTCCATCAAGAGGGCTCTTCCGACATCGCTCCACCCCTGCGAGCTGCCATC 360  
QY 427 TTTGCCAAGCAGAGGCTGCGCCCGAGAGCGGTTCTGTGGGGGGCATACTATCAGC 486  
Db 361 TTTGCCAAGCAGAGGCTGCGCCCGAGAGCGGTTCTGTGGGGGGCATACTATCAGC 420  
QY 487 TCCTGCTGGATTCTCTCTGCGGCCCATCTGCTTCCAGGAGAGGTTTCGCCCCCAACCTG 546

Db 421 TCCTGCTGGATTCTCTCTGCGGCCCACTGCTTCAGGAGAGGTTTCGCCCCCAACCTG 480  
QY 547 ACGTGATCTTGGCAGAACATACCGGTGTCTCCCTGCGAGGAGGAGCAGAAATTTGAA 606  
Db 481 ACGTGATCTTGGCAGAACATACCGGTGTCTCCCTGCGAGGAGGAGCAGAAATTTGAA 540  
QY 607 GTCGAAAAATACATTTGTCATAGGAATTCGATGATGACACTTACGACAAATGCAATTCG 666  
Db 541 GTCGAAAAATACATTTGTCATAGGAATTCGATGATGACACTTACGACAAATGCAATTCG 600  
QY 667 CTGCTGAGCTGAATTCGATTCCTCCCTGTCGCCAGAGAGCAGGTCGCGACT 726  
Db 601 CTGCTGAGCTGAATTCGATTCCTCCCTGTCGCCAGAGAGCAGGTCGCGACT 660  
QY 727 GTGTGCTTCCCGCGGACCTGCGAGTCCCGAGTGGAGTGTGAGCTCTCCGCG 786  
Db 661 GTGTGCTTCCCGCGGACCTGCGAGTCCCGAGTGGAGTGTGAGCTCTCCGCG 720  
QY 787 TACGCGAAGCATGAGGCTTGTCTCTCTTCTATTCGAGCGGCTGAAGAGGCTCATGTC 846  
Db 721 TACGCGAAGCATGAGGCTTGTCTCTCTTCTATTCGAGCGGCTGAAGAGGCTCATGTC 780  
QY 847 AGACTGTACCCATCCAGCCGCTGACATCAACATTTACTTAAACAGACAGTCAACGAC 906  
Db 781 AGACTGTACCCATCCAGCCGCTGACATCAACATTTACTTAAACAGACAGTCAACGAC 840  
QY 907 AACATGCTGTGCTGGAGACACTCGGAGCGGGGCCCCAGGCAAACTTGCACGACGCC 966  
Db 841 AACATGCTGTGCTGGAGACACTCGGAGCGGGGCCCCAGGCAAACTTGCACGACGCC 900  
QY 967 TGCAGGCGGATTCGGGAGGCCCCCTGCTGTGTCTGAACGATGCGCGCATGCTTTGTTG 1026  
Db 901 TGCAGGCGGATTCGGGAGGCCCCCTGCTGTGTCTGAACGATGCGCGCATGCTTTGTTG 960  
QY 1027 GGCATCATCAGCTGGGGCTTGGGCTGTGGAGAGGATGTCCCGGTGTGTACACAAAG 1086  
Db 961 GGCATCATCAGCTGGGGCTTGGGCTGTGGAGAGGATGTCCCGGTGTGTACACAAAG 1020  
QY 1087 GTTACCAACTACCTAGACTGGATTCTGTGACAAATGCGACCG 1128  
Db 1021 GTTACCAACTACCTAGACTGGATTCTGTGACAAATGCGACCG 1062

## RESULT 8

US-09-974-298-144  
; Sequence 144, Application US/09974298  
; Patent No. US20020156263A1  
; GENERAL INFORMATION:  
; APPLICANT: Chen, Hui-Mei  
; TITLE OF INVENTION: GENES EXPRESSED IN BREAST CANCER  
; FILE REFERENCE: PA-0037 P  
; CURRENT APPLICATION NUMBER: US/09/974,298  
; CURRENT FILING DATE: 2001-10-04  
; PRIOR APPLICATION NUMBER: 60/238,331  
; PRIOR FILING DATE: 2000-05-10  
; NUMBER OF SEQ ID NOS: 194  
; SOFTWARE: PERL Program  
; SEQ ID NO 144  
; LENGTH: 2641  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; OTHER INFORMATION: Incyte ID No. US20020156263A1 1001470CB1  
; NAME/KEY: unsure  
; LOCATION: 2635  
; OTHER INFORMATION: a, t, c, g, or other  
; OTHER INFORMATION: 144

Query Match 94.1%; Score 1061.8; DB 10; Length 2641;  
Best Local Similarity 99.3%; Pred. No. 0;  
Matches 1066; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 56 CCCAGCGCCCTCTGAGGAAACAGTACTGTCTTGGGAATGGGTGAGCTACCGTG 115  
Db 810 CCCCTGCGCTGCTGAGGAAACAGTACTGTCTTGGGAATGGGTGAGCTACCGTG 869  
QY 116 GCACGACAGCCTCACCGAGTCCGGTCTCTGCTCCGCTGGGAATCCATGATCCTGA 175  
Db 870 GCACGACAGCCTCACCGAGTCCGGTCTCTGCTCCGCTGGGAATCCATGATCCTGA 929  
QY 176 TAGGCAAGTTTACACAGACAGAACCCAGTCCCGAGCACTGGGCTTGGGCAACATA 235  
Db 930 TAGGCAAGTTTACACAGACAGAACCCAGTCCCGAGCACTGGGCTTGGGCAACATA 989  
QY 236 ATTACTCGCGAATCTGATGGGATGCCAAGCCTGTGTCACAGTGTGAAGACCGCA 295  
Db 990 ATTACTCGCGAATCTGATGGGATGCCAAGCCTGTGTCACAGTGTGAAGACCGCA 1049  
QY 296 GCGTCAGCGGAGTACTGTGATGTGCTCTGCTCCACCTGCGGCTGAGACAGTACA 355  
Db 1050 GCGTCAGCGGAGTACTGTGATGTGCTCTGCTCCACCTGCGGCTGAGACAGTACA 1109  
QY 356 GCCAGCCTCAGTTTCGCATCAAGAGGAGCTCTTCCCGACATCGCTCCCAACCCCTGGC 415  
Db 1110 GCCAGCCTCAGTTTCGCATCAAGAGGAGCTCTTCCCGACATCGCTCCCAACCCCTGGC 1169  
QY 416 AGGCTGCCATCTTCCAGACACAGAGTGGCCCGGAGAGCGTTCTGTGCGGGGCA 475  
Db 1170 AGGCTGCCATCTTCCAGACACAGAGTGGCCCGGAGAGCGTTCTGTGCGGGGCA 1229  
QY 476 TACTCATCAGCTCTGCTGATTTCTCTGCGCCGCACTGCTTCCAGGAGAGTTTCGC 535  
Db 1230 TACTCATCAGCTCTGCTGATTTCTCTGCGCCGCACTGCTTCCAGGAGAGTTTCGC 1289  
QY 536 CCCACACCTGACGCTGATTTGGGACAGAACATCCGGTGTCTCTGCGGAGGAGGAGC 595  
Db 1290 CCCACACCTGACGCTGATTTGGGACAGAACATCCGGTGTCTCTGCGGAGGAGGAGC 1349  
QY 596 AGAATTTGAAGTCGAAATACATTTCCATAGGAATTCGATGACACTTACGACA 1409  
Db 1350 AGAATTTGAAGTCGAAATACATTTCCATAGGAATTCGATGACACTTACGACA 1459  
QY 656 ATGACATTTGCGCTGCTGAGTGAATCGGATTCGCTGCTGCGGAGGAGGAGCAGG 715  
Db 1410 ATGACATTTGCGCTGCTGAGTGAATCGGATTCGCTGCTGCGGAGGAGGAGCAGG 1469  
QY 716 TGGTCCGACATGTGTGCTTCCCGCGGAGCTGACGCTGCGGAGTGGAGGAGTGTG 775  
Db 1470 TGGTCCGACATGTGTGCTTCCCGCGGAGCTGACGCTGCGGAGTGGAGGAGTGTG 1529  
QY 776 AGCTCTCGGCTACGGCAAGCATGAGGCTGTGCTCTTCTATTCGAGGCGGTGAAGG 835  
Db 1530 AGCTCTCGGCTACGGCAAGCATGAGGCTGTGCTCTTCTATTCGAGGCGGTGAAGG 1589  
QY 836 AGGCTCATGTGACAGTGTACCCATCCAGCGCTGACATCAACATTTACTTACAGAA 895  
Db 1590 AGGCTCATGTGACAGTGTACCCATCCAGCGCTGACATCAACATTTACTTACAGAA 1649  
QY 896 CAGTCAACGACAACTGTGTGTGCTGGAGACATCTCGAGGCGGCGGCGCCAGGCAAACT 955  
Db 1650 CAGTCAACGACAACTGTGTGTGCTGGAGACATCTCGAGGCGGCGGCGCCAGGCAAACT 1709  
QY 956 TGACAGAGCTGCGCAGGCGGATTCGGAGGCGGCTGTGTGTCTGAACGATGCGCGCA 1015  
Db 1710 TGACAGAGCTGCGCAGGCGGATTCGGAGGCGGCTGTGTGTCTGAACGATGCGCGCA 1769  
QY 1016 TGACTTTGTGGGCAATCATCAGCTGGGCTGGGCTGTGGAGAGGATGTCCGGGTG 1075  
Db 1770 TGACTTTGTGGGCAATCATCAGCTGGGCTGGGCTGTGGAGAGGATGTCCGGGTG 1829  
QY 1076 TGTACACAAAGTTTACCACTACCTAGACTGGATTCGTGACAAATGCGACCG 1128  
Db 1830 TGTACACAAAGTTTACCACTACCTAGACTGGATTCGTGACAAATGCGACCG 1882

RESULT 9  
US-10-007-926A-433  
; Sequence 433, Application US/10007926A  
; Publication No. US20030143539A1  
; GENERAL INFORMATION:  
; APPLICANT: BERTUCCI, FRANCOIS  
; APPLICANT: HOULGAITE, REMI  
; APPLICANT: BIRNBAUM, DANIEL  
; APPLICANT: NGUYEN, CATHERINE  
; APPLICANT: VIENS, PATRICE  
; APPLICANT: FERT, VINCENT  
; TITLE OF INVENTION: GENE EXPRESSION PROFILING OF PRIMARY BREAST CARCINOMAS  
; TITLE OF INVENTION: USING ARRAYS OF CANDIDATE GENES  
; FILE REFERENCE: 1546-R-00  
; CURRENT APPLICATION NUMBER: US/10/007.926A  
; CURRENT FILING DATE: 2001-12-07  
; PRIOR APPLICATION NUMBER: 60/254,090  
; PRIOR FILING DATE: 2000-12-08  
; NUMBER OF SEQ ID NOS: 468  
; SOFTWARE: Patent In Ver. 2.1  
; SEQ ID NO 433  
; LENGTH: 329  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: 5' terminal sequence. Homo sapiens plasminogen activator (PLAT)  
; NAME/KEY: modified\_base  
; LOCATION: (57)  
; OTHER INFORMATION: a, t, c or g  
; FEATURE:  
; NAME/KEY: modified\_base  
; LOCATION: (82)  
; OTHER INFORMATION: a, t, c or g  
; FEATURE:  
; NAME/KEY: modified\_base  
; LOCATION: (159)  
; OTHER INFORMATION: a, t, c or g  
; FEATURE:  
; NAME/KEY: modified\_base  
; LOCATION: (264)  
; OTHER INFORMATION: a, t, c or g  
; FEATURE:  
; NAME/KEY: modified\_base  
; LOCATION: (278)  
; OTHER INFORMATION: a, t, c or g  
US-10-007-926A-433  
  
Query Match 20.1%; Score 227.2; DB 12; Length 329;  
Best Local Similarity 97.4%; Pred. No. 5.9e-62;  
Matches 229; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
  
QY 894 AACAGTACCACGACAAATGCTGTCTGTGGAGACACTCGGAGCGGCGGCGCCAGGCAAA 953  
Db 1 AACAGTACCACGACAAATGCTGTCTGTGGAGACACTCGGAGCGGCGGCGCCAGGCAAA 60  
QY 954 CTTGCAGCAGCCCTGCGAGGCGGATTCGGAGGCGCCCTGTGTGTCTGAACGATGCGCG 1013  
Db 61 CTTGCAGCAGCCCTGCGAGGCGGATTCGGAGGCGCCCTGTGTGTCTGAACGATGCGCG 120  
QY 1014 CATGACTTTGGTGGGCAATCATCAGCTGGGCGCTGGGCTGTGGACAGAGGATGTCGCGG 1073  
Db 121 CATGACTTTGGTGGGCAATCATCAGCTGGGCGCTGGGCTNTGGACAGAGGATGTCGCGG 180  
QY 1074 TGTGTACAAAGGTTTACCACTACCTAGACTGGATTCGTGACAAATGCGACCG 1128  
Db 181 TGTGTACAAAGGTTTACCACTACCTAGACTGGATTCGTGACAAATGCGACCG 235  
  
RESULT 10  
US-09-880-503-15  
; Sequence 15, Application US/09880503  
; Patent No. US20020131964A1

QY	843	TGTCAGACTGTATCCCATCCAGCGCTGCGACATCACAAATTTACTTTAAACAGAACAGTCAC	902
Db	924	TGTGAAGCTGATTTTCCACCGGGAGTGTGAGCAGGCCCTACTACGGCTCTGAACTCAC	983
QY	903	CGACACATGCTGTGTGTCGGAGACACTCGGAGCGGGCCCCCAGGCAAACTTGACAGA	962
Db	984	CACCAAAATGCTATGTGCTGCTGACCCCC-----AATGGAAAACAGA	1025
QY	963	CGCTCCAGGGGGGATTCGGGAGGCCCCCTGTGTGCTCTGAACGATGGCCGCGATGACTTT	1022
Db	1026	TTCTTCCAGGGAGACTCAGGGGACCCCTGTGTCTTCCCTCAAGGCCGCGATGACTTT	1085
QY	1023	GGTGGGCATCATCAGCTGGGGCCCTGGGCTGTGTGACAGAAAGGATGTCCCGGGTGTGACAC	1082
Db	1086	GACTGGAAATTGTGAGCTGGGGCCGTGGATGTGCCCTGAAGGACAAGCCAGGGCTCTACAC	1145
QY	1083	AAAGTTACCACTACTAGACTGGATTGG	1112
Db	1146	GAGAGTCTCACACTTCTTACCTGGATCGG	1175

RESULT 11

US-10-198-846-12748

US-10-198-846-12748, Application US/10198846.

Sequence 12748, Application No. US2003009974A1

Publication No. US2003009974A1

GENERAL INFORMATION:

APPLICANT: Lillie, James

APPLICANT: Xu, Yongyao

APPLICANT: Wang, Youzhen

APPLICANT: Steinmann, Kathleen

TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS

TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND

TITLE OF INVENTION: THERAPY OF BREAST CANCER

FILE REFERENCE: MRI-049

CURRENT APPLICATION NUMBER: US/10/198,846

CURRENT FILING DATE: 2002-07-18

PRIOR APPLICATION NUMBER: 60/306,220

PRIOR FILING DATE: 2001-07-18

NUMBER OF SEQ ID NOS: 14084

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 12748

LENGTH: 1415

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: misc\_feature

LOCATION: 1, 2, 3, 1414, 1415

OTHER INFORMATION: n = A,T,C or G

US-10-198-846-12748

Query Match 16.7%; Score 188; DB 14; Length 1415;

Best Local Similarity 91.1%; Pred. No. 3.1e-49;

Matches 255; Conservative 0; Mismatches 20; Indels 5; Gaps 5;

QY	56	CCAGCGCGCTCTGAGGAAACAGTGACTGTCTACTTTGGGAATGGGTGAGCTACCGTG	115
Db	895	CCCTGCTGCTCTGAGGAAACAGTGACTG-TACTTTGGGAATGGGTGAGCTACCGTG	953
QY	116	GCAQGCACGCTCAACGAGTCGGGTGCTCTCT-GCTCCCGTGGAAATCCATGATCTCTG	174
Db	954	GCAQGCACAG-CTCACCGAGTCGGGTGCTCTCTGGCTCCCGTGGAAATCCATGATCTCTG	1012
QY	175	ATAGGCAAGTTTACACACACAGAACCCAGTCGCCAGGCACTGGGCTTGGGCAACAT	234
Db	1013	ATAGGCAAGTTTACACACACAGAACCCAGTCGCCAGGCACTGGGCTTGGGCAACAT	1072
QY	235	AATTACTGCGGAATCTCTGATGGGATGCAAGCCCTGTGTGACACGCTGTGAA-GAACCG	293
Db	1073	AAATTAATCTCCGAAT-CTGATGGGATGCAAGCCCTGTGTGACACGCTGTGTGAAACCG	1131
QY	294	CAGGCTGACGTGGGAGTACTGTGATGTGCCCTCTGCTCC	333

Db 1132 CAGGCTGACGTGGGAGTACTGTGATGTGGCCTCTGTGACC 1171

## RESULT 12

US-09-735-705-122

; Sequence 122, Application US/09735705

; Patent No. US20020052329A1

; GENERAL INFORMATION:

; APPLICANT: Wang, Tongtong

; APPLICANT: Fan, Liqun

; APPLICANT: Kalos, Michael D.

; APPLICANT: Bangor, Chaitanya S.

; APPLICANT: Hosken, Nancy

; APPLICANT: Fanger, Gary R.

; APPLICANT: Li, Samuel X.

; APPLICANT: Wang, Aijun

; APPLICANT: Skeiky, Yasir A.W.

; APPLICANT: Henderson, Robert A.

; APPLICANT: McNeill, Patricia D.

; APPLICANT: Fanger, Neil

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY

; FILE REFERENCE: 210121.455C14

; CURRENT APPLICATION NUMBER: US/09/735,705

; CURRENT FILING DATE: 2000-12-12

; NUMBER OF SEQ ID NOS: 419

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 122

; LENGTH: 1475

; TYPE: DNA

; ORGANISM: Homo sapien

US-09-735-705-122

Query Match

Best Local Similarity 15.1%; Score 170.8; DB 9; Length 1475;

Matches 551; Conservative 0; Mismatches 457; Indels 66; Gaps 4;

84 CTGCTACTTTGGGAATGGGTACGCTACCGTGGCAGCAGACGCTCACCAGTGGGTGC 143

287 CTGCTATGAGGGGAATGGTCACTTTTACCGAGGAAGCCAGCACTGACCCATGGGCG 346

144 CTCCTGCTCCGCTGGGAATTCATGCTGATAGGCAAGTTTACAGCACAGAACCC 203

347 GCCCTGCTCCGCTGGGAATTCGCACTGCTCCTCAGCAAAAGTACCATGCCACAGATC 406

204 CAGTGCACGACGCTGGGCTGGGCAACATAATTAATCTACCGGAATCTGATGGGATGC 263

407 TGATGCTTTTGGGCTGGGCTGGGGAACATAATTAATCTGAGGAACCCAGACACCGGAG 466

264 CAAAGCCCTGGTGCACGCTGCTGAAGAACCGCAGGCTGACGTGGGAGTACTGTGATGCC 323

467 GCGACCTGGTGTATGTGACAGTGGGCTTAAAGCCGCTTGTCCAAAGAGTGCATGGTGA 526

324 CTCCTGCTCCACCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGG 347

527 TGACTGCGCAGATGGAAGAAAGCCCTCTCTCTCCAGGAAGAAATTAATTAATTAATTA 586

348 ---ACAGTACAGCCAGCTCAGTTTCGATCAAGAGAGGCTCTTCGCGGACATCGCCTC 404

587 CCMAAGACTCTGAGGCGCGCTTTAAGATTTTGGGGAGAAATTCACACCATCGAGAA 646

405 CCACCCCTGGCAGGCTGCTATCTTTGCCAAGCAGAGAGGTCGCCCGGAGAGCGGTTCTCT 464

647 CCAGCCCTGGTTTGGGCTGCTACAGAGGACCGCGGG---GGGCTCTGTCACTACGT 703

465 GTGCGGGGCTATCATCATGCTCTGCTGATTTCTCTGCGGCGGCTCTGCTGCTGCTGCTG 524

704 GTGTGGAGGCGCTCATCAGCCCTTGTGCTGGTGTATCAGCGCCACACACTGCTTCTTGA 763

525 GAGGTTTCCGCGCCACCACTGACGCTGATCTTGGGAGAAATACACCGGCTGGTCCCTGG 584

764 TTACCAAGAGGAGGACTACATCTGCTTACCTGGTCTGCTCAAGGCTTAATCTCAACAC 823

QY 585 CGAGGAGGAGCAAAATTTGAAGTCGAAATAATACATTTGTCATTAAGGAATTCGATGATGA 644

Db 824 GCAAGGGGAGATGAAGTTTGAAGTGGAAACCTCTCTACACAGGACTACAGCGCTGA 883

QY 645 CA-----CTTACGACAATGACATTTGGCTGCTGAGCTGAAATCGGATTCGTCGCGTG 698

Db 884 CACGCTTGTCTACACACAGCATTTGCTTGAAGATCCGTTTCAAGAGGAGGAGGTG 943

QY 699 TGCCGAGGAGGAGCGGTGTCGCTTCCGCTGCTGCTTCCCGGCGGACCTGCGAGTCCC 758

Db 944 TGCGCAGCATCCCGGACTATACAGACCATCTGCTGCTCGATGTATAACGATCCCA 1003

QY 759 GGACTGAGGAGGTGCTGAGCTCTCCGCTACCGCAAGCATGAGGCTTGTCTCTTCTA 818

Db 1004 GTTTGGCACAGCTGTGAGATCACTGGCTTTGGAAAGAGAAATTTACCGCACTATCTCTA 1063

QY 819 TTGCGAGGCTGAAGGAGGCTCATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 878

Db 1064 TCCGAGCAGCTGAAGATGACTGTTGAGCTGATTTCCACCGGAGTGTACGAGCC 1123

QY 879 ACATTTACTTAACAGAAACAGTCAACGACATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 938

Db 1124 CCACTACTACGCTCTGAGTCAACCAACCAATCTGCTGCTGCTGCTGCTGCTGCTGCTG 1180

QY 939 CGGCCCCAGCAAACTTTGACAGGCTGCGGAGGCTGCGGAGGCTGCGGAGGCTGCGGAG 998

Db 1181 -----GAAACAGATTTCTGCGAGGAGACTCAGGGGAGCCCTCGCTCTG 1225

QY 999 TCTGAAGATGGCGCATGACTTTGCTGGGCTCATCAGCTGCGGCTGCGGCTGCGGCTGCG 1058

Db 1226 TTCCCTCAAGGCGCGCATGACTTTGCTGGAATTTGAGCTGGGCGGCTGATGTCCTT 1285

QY 1059 GAAGGATGCTCCGCTGCTGTACACAAAGGTTTACCAACTTACCTAGACTGGATTCG 1112

Db 1286 GAAGGACAGCCAGCGCTTACACAGAGTCTCACACTTCTTACCTGGATCCG 1339

## RESULT 13

US-09-850-716A-122

; Sequence 122, Application US/09850716A

; Patent No. US20020115139A1

; GENERAL INFORMATION:

; APPLICANT: Kalos, Michael D.

; APPLICANT: McNeill, Patricia D.

; APPLICANT: Retter, Marc W.

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY

; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER

; FILE REFERENCE: 210121.455C15

; CURRENT APPLICATION NUMBER: US/09/850,716A

; CURRENT FILING DATE: 2001-05-07

; NUMBER OF SEQ ID NOS: 440

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 122

; LENGTH: 1475

; TYPE: DNA

; ORGANISM: Homo sapien

US-09-850-716A-122

Query Match

Best Local Similarity 15.1%; Score 170.8; DB 10; Length 1475;

Matches 551; Conservative 0; Mismatches 457; Indels 66; Gaps 4;

84 CTGCTACTTTGGGAATGGGTACGCTACCGTGGCAGCAGCCTCACCGAGTGGGTGC 143

287 CTGCTATGAGGGGAATGGTCACTTTTACCGAGGAAGCCAGCACTGACACCATGGGCG 346

144 CTCCTGCTCCGCTGGGAATTCATGATCTCTGATAGGCAAGTTTACAGCACAGAACCC 203

347 GCCCTGCTCCGCTGGGAATTCGCACTGCTCTTCCAGCAAAAGTACCATGCCACAGATC 406

204 CAGTGCACGACATGGGCTGGGCAACATAATTAATCTGCGGAAATCTGATGGGATGC 263

407 TGATGCTCTTTCAGCTGGGCTGGGGAACATAATTAATCTGATGAGGAACCCAGAAC 466

```
QY 264 CAAGCCCTGGTCCAGCTGCTGAAGAACCGCAGGCTGACGTGGGAGTACTGTGATGTGCC 323
Db |||||
QY 467 GCGACCTGCTGTATGTGACAGTGGGCTTAAGCGCTTGTCCAGAGTGTGATGTGCA 526
Db |||||
QY 324 CTCCTGCTCCACTCGCGGCTGAG----- 347
Db |||||
QY 527 TGACTGCGCAGATGGAAGAAAGCCCTCTCTCTCCAGAGAATTAATAATTCAGTGTGG 586
Db |||||
QY 348 ---ACAGTACAGCAGCTCAGTTCGATCAAGAGGCTTTCGCGGACATCGCTC 404
Db |||||
QY 587 CCAAGAAGACTGTGAGCGCCGCTTTAAGATTATGGGGAGAAATTCACCAACATCGAAA 646
Db |||||
QY 405 CCACCCCTGGCAGGCTGCCATCTTTGCCAAGCACAGGAGTGGCCCGGAGAGCGGTCTCT 464
Db |||||
QY 647 CCAGCCCTGTTTGGCGCATCTACAGGAGGACCGGG---GGGCTCTGTCACTAGT 703
Db |||||
QY 465 GTGGGGGCGATCATCATCAGCTCTGTGTGATTTCTCTGCGCGCCACTGTCTCCAGGA 524
Db |||||
QY 704 GTGTGAGGAGCGCTCATCAGCCCTTGTGGGTGATCAGCGGCACACACTGCTTCATTGA 763
Db |||||
QY 525 GAGGTTTCGCGCCCACTCAGCGTGTATCTTGGCAGAGATCGCCCGGAGATCCCGTGG 584
Db |||||
QY 764 TTACCCAAAGAGAGGAGTACTACCTGTCTACCTGGGTGCTCAAGGCTTAACCTCAACAC 823
Db |||||
QY 585 CGAGGAGGAGCAGAAATTTGAAGTTCGAAATATACATTGTCCATAAGGAATTCGATGTA 644
Db |||||
QY 824 GCAAGGGAGATGAATTTGAGTGGAAACCTCATCTTACAGAGGACTACAGGCTGA 883
Db |||||
QY 645 CA-----CTTACGACAATGACATTCGCTGCTGAGCTGAAATCGGATTCGTCCCGCTG 698
Db |||||
QY 884 CACGCTTGTCTACCAACAGCATTTGCTTGTGAAGATCCGTTTCCAAGGAGGCGAGTG 943
Db |||||
QY 699 TGCCGAGGAGCAGCGTGGTCCGACTGTGTGGCTTCCCGCGGAGCTGACGCTGCC 758
Db |||||
QY 944 TGCGCAGCCATCCCGGACTATACAGACCATCTGCCTGCCCTCGATGTATAAGATCCCCA 1003
Db |||||
QY 759 GAGCTGAGCGAGTGTGAGCTCTCGGCTTACGGCAAGCATGAGGCTTGTCTCTTTCTA 818
Db |||||
QY 1004 GTTTCGCACAGCTGTGAGATCACTGGCTTTGGAAGAGAAATTCACGAGTATCTTA 1063
Db |||||
QY 819 TTCGAGCGGCTGAAGAGGCTGATGTGAGTGTGATGATGATGATGATGATGATGATGAT 878
Db |||||
QY 1064 TCCGAGCAGCTGAAGATGACTTGTGTGAAGTGTATTTCCACCGGAGTGTGACGAGCC 1123
Db |||||
QY 879 ACATTTACTTAACAGAACGTCACCGACAAATGTGTGTGTGTGTGTGTGTGTGTGTGTGT 938
Db |||||
QY 1124 CCACCTACTACGGCTCTGAAGTCAACCAACCAAAATGCTGTGTGTGTGTGTGTGTGTGTGT 1180
Db |||||
QY 939 CGGGCCCCAGGCAAACTTGACAGCGCTTCCAGGCGCATTCGGGAGGCCCCCTGGTGTG 998
Db |||||
QY 1181 -----GAAACAGATTCCTGCCAGGAGACTCAGGGGAGCCCCCTGCTGTG 1225
Db |||||
QY 999 TCTGAACGATGCGCATGACTTTGGTGGGATCATCAGCTGGGCGCTGGGCTGTGGACA 1058
Db |||||
QY 1226 TTCCCTCCAAGCGCGCATGACTTTGACTGGAATTTGAGCTGGGCGCTGGATGTGCCCT 1285
Db |||||
QY 1059 GAAGATGTCGGGTGTGTACACAAAGTTACCAACTACCTAGACTGGATTGCG 1112
Db |||||
QY 1286 GAAGGACAAGCGGCGCTTACACAGAGTCTCACACTTCTTACCCCTGGATCCG 1339
Db |||||
```

## RESULT 14

US-09-987-778-122

; Sequence 122, Application US/09897778

; Patent No. US20020147143A1

; GENERAL INFORMATION:

; APPLICANT: Wang, Tongtong

; APPLICANT: Marnerakis, Margarita

; APPLICANT: Fanger, Gary R.

; APPLICANT: Vedwick, Thomas S.

; APPLICANT: Carter, Derrick

; APPLICANT: Watanabe, Yoshihiro

```
; APPLICANT: Henderson, Robert A.
; APPLICANT: Peckham, David W.
; APPLICANT: Fanger, Neil
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.455C16
; CURRENT APPLICATION NUMBER: US/09/897,778
; NUMBER OF SEQ ID NOS: 467
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 122
; LENGTH: 1475
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-987-778-122
```

## Query Match

15.1%; Score 170.8; DB 10; Length 1475;

Best Local Similarity 51.3%; Pred. No. 9.6e-44;

Matches 551; Conservative 0; Mismatches 457; Indels 66; Gaps 4;

```
QY 84 CTGCTACTTTGGGAATGGGTGAGCTTACCGTGGCAGCAGCAGCTCCAGGATCGGGTGC 143
Db |||||
QY 287 CTGCTATGAGGGGAATGGTCACTTTTACCGAGGAAAGCCAGCACTGACACCATGGGCGC 346
Db |||||
QY 144 CTCCTGCTCCCGTGGAAATCCATGATCCTGATAGGCAAGGTTTACAGCAGACAGAACCC 203
Db |||||
QY 347 GCGCTGCTGCGCTGGAACTGTGCCACTGTCTTCAGCAACGTAACATGCCACAGATC 406
Db |||||
QY 204 CAGTGGCCAGGCACTGGGCTGGGCAACATAATTAATGCGGGAATCTGTATGGGGATGC 263
Db |||||
QY 407 TGATGCTCTTTCAGCTGGGCTGGGAAACATAATTAATGCGAGGAACCCAGACCAACCGGAG 466
Db |||||
QY 264 CAAGCCCTGGTGGCAGCTGCTGAAGAACCGCAGGCTGACGTGGGAGTACTGTGATGTGCC 323
Db |||||
QY 467 GCGACCTGTGTGTATGTGAGGTGGGCTTAAAGCCGCTTGTCCAAGAGTGCATGGTGCA 526
Db |||||
QY 324 CTCCTGCTCCACCTGCGGCGCTGAG----- 347
Db |||||
QY 527 TGACTGCGCAGATGGAAGAAAGCCCTCTCTCTCCAGAGAATTAATAATTCAGTGTGG 586
Db |||||
QY 348 ---ACAGTACAGCAGCTCAGTTCGATCAAGAGGAGGCTTTCGCGGACATCGGCTC 404
Db |||||
QY 587 CCAAGAGACTCTGAGGCGCGCTTTAAGATTATTTGGGGAGAAATTCACCACTCGAGAA 646
Db |||||
QY 405 CCACCCCTGGCAGGCTGCCATCTTTGCCAAGCACAGGAGTGGCCCGGAGAGCGGTCTCT 464
Db |||||
QY 647 CCAGCCCTGTTTGGCGCATCTACAGGAGGACTACATCGTCTACCTGGGTGCTTAACCTCAACAC 823
Db |||||
QY 465 GTGGGGGCGATACTCATCAGCTCTCTGATGATTTCTCTGCGCGCCACTGTCTCCAGGA 524
Db |||||
QY 704 GTGTGAGGAGCGCTCATCAGCCCTTGTGGGTGATCAGCGGCACACACTGCTTCATTGA 763
Db |||||
QY 525 GAGTTCGCGCCCACTCAGCGTGTATCTTGGGAGAAATCATTAAGAAATTCGATGATGA 584
Db |||||
QY 764 TTACCCAAAGAGAGGAGTACTACATCGTCTACCTGGGTGCTCAAGGCTTAACCTCAACAC 823
Db |||||
QY 585 CGAGGAGGAGCAGAAATTTGAAGTTCGAAATATACATTGTCCATAAGGAATTCGATGTA 644
Db |||||
QY 824 GCAAGGGAGATGAATTTGAGTGGAAACCTCATCTTACAGAGGACTACAGGCTGA 883
Db |||||
QY 645 CA-----CTTACGACAATGACATTCGCTGCTGAGCTGAAATCGGATTCGTCTCTTTCTA 818
Db |||||
QY 1004 GTTTCGCACAGCTGTGAGATCACTGGCTTTGGAAGAGAAATTCACGAGTATCTTA 1063
Db |||||
QY 819 TTCGAGCGGCTGAAGAGGCTGATGTGAGTGTGATGATGATGATGATGATGATGATGAT 878
Db |||||
QY 1064 TCCGAGCAGCTGAAGATGACTTGTGTGAAGTGTATTTCCACCGGAGTGTGACGAGCC 1123
Db |||||
QY 879 ACATTTACTTAACAGAACGTCACCGACAAATGTGTGTGTGTGTGTGTGTGTGTGTGTGT 938
Db |||||
QY 1124 CCACCTACTACGGCTCTGAAGTCAACCAACCAAAATGCTGTGTGTGTGTGTGTGTGTGTGT 1180
Db |||||
QY 939 CGGGCCCCAGGCAAACTTGACAGCGCTTCCAGGCGCATTCGGGAGGCCCCCTGGTGTG 998
Db |||||
QY 1181 -----GAAACAGATTCCTGCCAGGAGACTCAGGGGAGCCCCCTGCTGTG 1225
Db |||||
QY 999 TCTGAACGATGCGCATGACTTTGGTGGGATCATCAGCTGGGCGCTGGGCTGTGGACA 1058
Db |||||
QY 1226 TTCCCTCCAAGCGCGCATGACTTTGACTGGAATTTGAGCTGGGCGCTGGATGTGCCCT 1285
Db |||||
QY 1059 GAAGATGTCGGGTGTGTACACAAAGTTACCAACTACCTAGACTGGATTGCG 1112
Db |||||
QY 1286 GAAGGACAAGCGGCGCTTACACAGAGTCTCACACTTCTTACCCCTGGATCCG 1339
Db |||||
```



**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 18, 2003, 22:44:33 ; Search time 3299.95 Seconds  
(without alignments)  
8307.845 Million cell updates/sec

Title: US-09-987-455-2

Perfect score: 1128  
Sequence: 1 atgaaaagacagctatgc.....ttcgtgacaacatgcgacgcg 1128

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

EST:\*

1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hrc:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hrc:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: em\_gss\_hum:\*  
18: em\_gss\_inv:\*  
19: em\_gss\_pin:\*  
20: em\_gss\_vrt:\*  
21: em\_gss\_fun:\*  
22: em\_gss\_mam:\*  
23: em\_gss\_mus:\*  
24: em\_gss\_pro:\*  
25: em\_gss\_rod:\*  
26: em\_gss\_phg:\*  
27: em\_gss\_vrl:\*  
28: gb\_gss1:\*  
29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
c 1	895	79.3	930	13	BX329047
2	853	75.6	912	13	BU149958
3	850.2	75.4	962	13	BX389609
4	818.6	72.6	891	13	BU146167

5	813.8	72.1	993	13	BQ278177
6	811.2	71.9	883	13	BQ690978
7	807.6	71.6	918	13	BX389608
8	796.2	70.6	930	13	BU845014
9	782.8	69.4	888	13	BU157720
10	785.2	67.8	950	13	BQ927899
11	784.4	67.8	916	12	BI765247
12	706.8	62.7	948	10	BE616613
13	700.4	62.1	958	12	BG763582
14	694.4	61.6	974	13	BQ960216
15	693.4	61.5	709	10	BE732704
16	688.2	61.0	1060	12	BM556999
17	677	60.0	1042	10	BE732413
18	672	59.6	957	13	BQ687538
19	665.8	59.0	948	13	BQ687779
20	665.2	59.0	704	9	AU134301
21	655.4	58.1	818	9	AU124602
22	655.2	58.1	903	13	BQ689095
23	653.2	57.9	830	13	BQ896938
24	653.2	57.9	867	13	BU179903
25	652.2	57.8	904	13	BQ689840
26	652.2	57.8	911	13	BQ684734
27	651	57.7	818	12	BM006474
28	649.6	57.6	797	13	BQ879911
29	642.2	56.9	1037	13	BU157433
30	631.2	56.0	813	10	BG120780
31	629.8	55.8	663	9	AU134463
32	624	55.3	624	10	BE549489
33	585.4	51.9	587	10	BE394944
34	561	49.7	561	9	AL047284
35	553.4	49.1	555	14	CB128855
36	535.4	47.5	537	14	CB155052
37	534	46.1	964	12	BG769384
38	520.4	46.1	905	13	BU859271
39	514.8	45.6	883	14	CD359959
40	514.4	45.6	516	14	CB140923
41	493.4	43.7	916	13	BU157241
42	487.8	43.2	527	9	AW630119
43	482.8	42.8	553	9	AW394127
44	474.8	42.1	917	13	BU190904
45	472.8	41.9	937	13	BQ679501

#### ALIGNMENTS

RESULT 1  
BX329047/c  
LOCUS BX329047 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens CDNA  
DEFINITION clone CSODI039YK15 3-PRIME, mRNA sequence.  
ACCESSION BX329047  
VERSION BX329047.1 GI:30307809  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 930)  
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.  
Full-length cDNA libraries and normalization  
Unpublished  
Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 EVRY cedex - France  
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 4498.r For more information about this cluster, see  
http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CS08AI021ZD08\_C  
S019801\_cluster=4498.r. Contact : Feng Liang Email :  
fliang@lifetech.com URL : http://fulllength.invitrogen.com/  
Invitrogen Corporation 1600 Faraday Avenue Genoscope sequence ID :

```
CS0BA1021ZD08_CS01980.1.
FEATURES
  Location/Qualifiers
    1..930
      /organism="Homo sapiens"
      /mol_type="mRNA"
      /db_xref="taxon:9606"
      /clone="CS0D1039YK15"
      /tissue_type="PLACENTA COT 25-NORMALIZED"
      /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
      /note="1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT      184 a   268 c   282 g   194 t
ORIGIN
Query Match      79.3%; Score 895; DB 13; Length 930;
Best Local Similarity 98.6%; Pred. No. 2.9e-222;
Matches 912; Conservative 0; Mismatches 12; Indels 1; Gaps 1;
QY      67  TCTGAGGGAACAGTACTGCTACTTTGGGAATGGTCCAGCTACCGTGGGACGACACAGC 126
Db      927  TTGAGGGAACAGTACTGCTACTTGGGAATGGTCCAGCTACCGTGGGACGACACAGC 869
QY      127  CTCACCGAGTCGGGTGCTCTCTGCTCCCGTGGAAATTCATGATCCTGATAGGCAAGGTT 186
Db      868  CTCACCGAGTCGGGTGCTCTCTGCTCCCGTGGAAATTCATGATCCTGATAGGCAAGGTT 809
QY      187  TACACAGACAGAACCCAGTCCCGAGGACCTGGGCTGGGCAACATTAATTAATCTGCCGG 246
Db      808  TACACAGACAGAACCCAGTCCCGAGGACCTGGGCTGGGCAACATTAATTAATCTGCCGG 749
QY      247  AATCTGTATGGGATGCCAAGCCCTGTGTGCCACGCTGTGAGAACCGCAGGCTGACGTGG 306
Db      748  AATCTGTATGGGATGCCAAGCCCTGTGTGCCACGCTGTGAGAACCGCAGGCTGACGTGG 689
QY      307  GAGTACTGTGATGTCCCTCTCTGCTCCACCTCGGCTGAGACAGTACAGCAGGCTCAG 366
Db      688  GAGTACTGTGATGTCCCTCTCTGCTCCACCTCGGCTGAGACAGTACAGCAGGCTCAG 629
QY      367  TTTGCGATCAAGAGGAGGCTTTTCCGCGACATCGCTCCACCCCTGGAGAGGCTGCCATC 426
Db      628  TTTGCGATCAAGAGGAGGCTTTTCCGCGACATCGCTCCACCCCTGGAGAGGCTGCCATC 569
QY      427  TTTGCCAAGCAGCAGAGTTCGCCGAGAGCGGTTCTGTGCGGGGATCTCATATCAGC 486
Db      568  TTTGCCAAGCAGCAGAGTTCGCCGAGAGCGGTTCTGTGCGGGGATCTCATATCAGC 509
QY      487  TCCTGCTGGATTTCTCTCTGCCGCCCACTGCTTCCAGGAGAGGTTTCCGCCCCACCCACTG 546
Db      508  TCCTGCTGGATTTCTCTCTGCCGCCCACTGCTTCCAGGAGAGGTTTCCGCCCCACCCACTG 449
QY      547  ACGGTGATCTTGGGAGAACATACCGGGTGGTTCCTGGCGAGGAGGACAGAAATTTGAA 606
Db      448  ACGGTGATCTTGGGAGAACATACCGGGTGGTTCCTGGCGAGGAGGACAGAAATTTGAA 389
QY      607  GTGGAATAATCATTTGCTCAATGAATTCGATGATGACATTACGACAAATGACATTCGG 666
Db      388  GTGGAATAATCATTTGCTCAATGAATTCGATGATGACATTACGACAAATGACATTCGG 329
QY      667  CTGCTGCGAGCTGAATTCGCTCCGCTGTGCTCCAGGAGAGCAGCGTGTCTCGCACT 726
Db      328  CTGCTGCGAGCTGAATTCGCTCCGCTGTGCTCCAGGAGAGCAGCGTGTCTCGCACT 269
QY      727  GTGTGCTTCCCGCGGACCTGAGCTGCCGAGCTGGAGAGGATGTGAGCTCTCCGGC 786
Db      268  GTGTGCTTCCCGCGGACCTGAGCTGCCGAGCTGGAGAGGATGTGAGCTCTCCGGC 209
QY      787  TACGCAAGCATGAGGCTTGTCTCTTCTTATTCGGAGCGGCTGAAGAGGCTCATGTC 846
Db      208  TACGCAAGCATGAGGCTTGTCTCTTCTTATTCGGAGCGGCTGAAGAGGCTCATGTC 149
QY      847  AGACTGTACCCATCCAGCGGCTGCACATCAACAATTTTACTTAACAGAACAGTCAACGAC 906
148  AGACTGTACCCATCCAGCGCTGCACATCAACAATTTTACTTAACAGAACAGTCAACGAC 89
907  AACATGCTGTGTCTGTGAGACACTCGGAGCGGGGGCCCGCAAACTTTCACGACGCC 966
88  AACATGCTGTGTGTGTGAGACACTCGGAGCGGGGGCCCGCAAACTTTCACGACGCC 29
967  TGCCAGGCGGATTCGGGAGGCCGCC 991
28  TGCCAGGCGGATTCGGGAGGCCGCC 4
RESULT 2
LOCUS      BUI149958      912 bp      mRNA      linear      EST 03-SEP-2002
DEFINITION AGENCOURT 8137213 Lupski_dorsal_root_ganglion Homo sapiens cDNA
ACCESSION   BUI149958
VERSION     BUI149958.1 GI:22663490
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
REFERENCE   1 (bases 1 to 912)
AUTHORS    NIH-MGC http://mgc.nci.nih.gov/.
TITLE      National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL     Unpublished
COMMENT     Contact: Robert Strauberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. James R. Lupski
            CDNA Library Preparation: Life Technologies, Inc.
            DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone Distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM13572 row: m column: 16
            High quality sequence start: 18
            High quality sequence stop: 644.
FEATURES
  Location/Qualifiers
    1..912
      /organism="Homo sapiens"
      /mol_type="mRNA"
      /db_xref="taxon:9606"
      /clone="IMAGE:6184119"
      /sex="male"
      /tissue_type="dorsal root ganglia"
      /dev_stage="adult, 36 yr"
      /lab_host="DH10B"
      /clone_lib="Lupski dorsal root ganglion"
      /note="vector: pCMV-SPORT6 (Life Technologies); Site_1:
            NotI; Site 2: SalI; cDNA made by oligo-dT priming.
            Directionally cloned using the following adaptors:
            5'-TCGACCCAGCGCTCCG-3' and
            5'-GACTAGTTCGTAGTCGAGCGCGCCCT(15)-3'. Size selected >
            1 kb for average insert length 1.7 kb. This is a primary
            library, non-amplified. Library constructed by Life
            Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
            College of Medicine) and is available through Life
            Technologies."
BASE COUNT      189 a   272 c   271 g   179 t
ORIGIN
Query Match      75.6%; Score 853; DB 13; Length 912;
Best Local Similarity 98.9%; Pred. No. 2.5e-211;
Matches 901; Conservative 0; Mismatches 5; Indels 5; Gaps 4;
QY      197  AGAACCCAGTCCCGAGGCTGGGCTGGGGCTGGGGCTGGGGCTGGGGCTGGGGCTGGG 254
Db      1  AGAACCCAGTCCCGAGGCTGGGCTGGGGCTGGGGCTGGGGCTGGGGCTGGGGCTGGG 60
255  TGGGGATGCCAAGCCCTGGTGCACCGTGTCTGGAAGAACCGCAGGCTGACGTGGGAGTACTG 314
```

Db 61 TGGGATGCCAAGCCCTGGTCCAGTCTGAAGAACCAGCGCTGACGTGGGAGTACTG 120  
Qy 315 TGATGTGCCCTCTCTCCACCTGGCGCTGAGACAGTACAGCCAGCCTCAGTTTCGCAT 374  
Db 121 TGATGTGCCCTCTCTCCACCTGGCGCTGAGACAGTACAGCCAGCCTCAGTTTCGCAT 180  
Qy 375 CAAAGGAGGCTCTTTCGGCGCATCGCTCCACCCCTGGAGGCTGCATCTTTGCCAA 434  
Db 181 CAAAGGAGGCTCTTTCGGCGCATCGCTCCACCCCTGGAGGCTGCATCTTTGCCAA 240  
Qy 435 GCACAGGAGTCCGCCGAGAGCGGTCTCTGTGGGGGCATCTCATCAGCTCTCTGCTG 494  
Db 241 GCACAGGAGTCCGCCGAGAGCGGTCTCTGTGGGGGCATCTCATCAGCTCTCTGCTG 300  
Qy 495 GATTCTCTCTGCGGCCACTGCTTCCAGGAGAGGTTTCCGCCCAACCACTGACGGTGTAT 554  
Db 301 GATTCTCTCTGCGGCCACTGCTTCCAGGAGAGGTTTCCGCCCAACCACTGACGGTGTAT 360  
Qy 555 CTTGGGCGAACAATACCGGGTGTCTCTGGCGAGGAGGAGGAGAAATTTGAAGTCGAAAA 614  
Db 361 CTTGGGCGAACAATACCGGGTGTCTCTGGCGAGGAGGAGGAGAAATTTGAAGTCGAAAA 420  
Qy 615 ATACATTCTCCATAGGAATTCGATGATGACACTTACGACATGACATTCGCTGCTGCA 674  
Db 421 ATACATTCTCCATAGGAATTCGATGATGACACTTACGACATGACATTCGCTGCTGCA 480  
Qy 675 GCTGAAATCGAATTCGCTCCGCTGTGCCAGAGAGCGGTGCTCCCACTGTGTGCT 734  
Db 481 GCTGAAATCGAATTCGCTCCGCTGTGCCAGAGAGCGGTGCTCCCACTGTGTGCT 540  
Qy 735 TCCCCCGCGGACCTGCGAGCTGCCGACTGGAAGGAGTGTGAGCTCTCCGCTACGGCAA 794  
Db 541 TCCCCCGCGGACCTGCGAGCTGCCGACTGGAAGGAGTGTGAGCTCTCCGCTACGGCAA 600  
Qy 795 GCATGAGGCTTGTCTCTTCTTCTATTCGAGGCGCTGGAAGGCTCATGTCACTGTA 854  
Db 601 GCATGAGGCTTGTCTCTTCTTCTATTCGAGGCGCTGGAAGGCTCATGTCACTGTA 660  
Qy 855 CCATCCAGCGCTGCATCACACATTTACTTAAAGAACAGTCCACGACACATGCT 914  
Db 661 CCATCCAGCGCTGCATCACACATTTACTTAAAGAACAGTCCACGACACATGCT 720  
Qy 915 GTGTCTCGAGACACTCGAGCGGGCGGCCAGCAAACTTGCACAGCGCTGCGAGG 974  
Db 721 GTGTCTCGAGACACTCGAGCGGGCGGCCAGCAAACTTGCACAGCGCTGCGAGG 780  
Qy 975 -CGATTCCGAGG-CCCCCTGTGTCTGAACAGATGCGCGCATGCTTTGGTGGGATC 1032  
Db 781 NCGATTCCGAGGCGGCCCTGTGTCTGAACAGATGCGCGCATGCTTTGGTGGGATC 840  
Qy 1033 ATCAGTGGGCGCTGGGCTGTGGACAGAGGATGT-CCCCGGTGTGTACAAAGGTTAC 1091  
Db 841 ATCAGTGGGCGCTGGGCTGTGGACAGAGGATGTCCCCGGTGTGTACAAAGGTTAC 900  
Qy 1092 CAACTACCTAG 1102  
Db 901 AAATACCTAG 911

RESULT 3  
LOCUS BX389609  
DEFINITION BX389609 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA  
clone CS0D1039YK15 5-PRIME, mRNA sequence.  
ACCESSION BX389609  
VERSION BX389609.1 GI:30463470  
KEYWORDS EST.  
SOURCE Homo sapiens  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 962)

AUTHORS  
TITLE  
JOURNAL  
COMMENT

Li, W.B., Gruber, C., Jessee, J. and Polayes, D.  
Full-length cDNA libraries and normalization  
Unpublished  
Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 EVRY cedex - France  
Email: segrif@genoscope.cns.fr, Web : www.genoscope.cns.fr  
was normalized. Library was constructed by Life Technologies, a  
division of Invitrogen. This sequence belongs to sequence cluster  
4498.r For more information about this cluster, see  
http://www.genoscope.cns.fr/  
cgi-bin/cluster.cgi?seq=CS0BAF025ZG09\_AF02376\_3&cluster=4498.r.  
Contact : Feng liang Email : fliang@lifetech.com URL :  
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600  
Paradise Avenue Genoscope sequence ID : CS0BAF025ZG09\_AF02376\_3.

FEATURES  
source

1..962  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CS0D1039YK15"  
/tissue\_type="PLACENTA COT 25-NORMALIZED"  
/clone\_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"  
/note="1st strand cDNA was primed with a NotI-oligo (dT)  
primer. Five prime end enriched, double-strand cDNA was  
digested with NotI and cloned into the NotI and EcoRV  
sites of the pCMVSPORT 6 vector. Library was normalized."  
BASE COUNT 203 a 290 c 277 g 191 t 1 others

Query Match 75.4%; Score 850.2; DB 13; Length 962;  
Best Local Similarity 96.6%; Pred. No. 1.4e-210;  
Matches 879; Conservative 0; Mismatches 29; Indels 2; Gaps 1;

Qy 208 GCCCAGGACATGGCCCTGGCGCAACATAATTAATCTGCGGAACTCTGTATGGGATGCCAAG 267  
Db 1 GCCCAGGACATGGNGCCTGGGCAAACTAATTAATCTGCGGAACTCTGTATGGGATGCCAAG 60  
Qy 268 CCCTGGTGCCACGTGCTGAAGAACCAGCGGTGACGTGGGAGTACTGTGATGTGCCCTCC 327  
Db 61 CCCTGGTGCCACGTGCTGAAGAACCAGCGGTGACGTGGGAGTACTGTGATGTGCCCTCC 120  
Qy 328 TGCTTCCACCTGCGGCTTGACAGTACAGCAGCCTCAGTTTCGCATCAAAGGAGGCTC 387  
Db 121 TGCTTCCACCTGCGGCTTGACAGTACAGCAGCCTCAGTTTCGCATCAAAGGAGGCTC 180  
Qy 388 TTGCGCCAGATCGGCTCCGACCCCTGGCAGGCTGCCATCTTGCACAGCAGAGGCTCG 447  
Db 181 TTGCGCCAGATCGGCTCCGACCCCTGGCAGGCTGCCATCTTGCACAGCAGAGGCTCG 240  
Qy 448 CCGGAGAGCGGTTCCTGTGCGGGGCACTACTCATCAGCTCCTGCTGGATTCTCTCTGCC 507  
Db 241 CCGGAGAGCGGTTCCTGTGCGGGGCACTACTCATCAGCTCCTGCTGGATTCTCTCTGCC 300  
Qy 508 GCCCAGCTTCCAGGAGAGGTTTCCGCCCAACCACTGACGCTGATCTTTGGGCGAACA 567  
Db 301 GCCCAGCTTCCAGGAGAGGTTTCCGCCCAACCACTGACGCTGATCTTTGGGCGAACA 360  
Qy 568 TACCGGCTGCTCCCTGGCGAGGAGGAGCAAAATTTGAAGTCGAAAATACATTGTCCAT 627  
Db 361 TACCGGCTGCTCCCTGGCGAGGAGGAGCAAAATTTGAAGTCGAAAATACATTGTCCAT 420  
Qy 628 AAGGAATTCGATGATGACACTTACGACAAATGACATTTCGCTGCTGCTGAGCTGAAATCGGAT 687  
Db 421 AAGGAATTCGATGATGACACTTACGACAAATGACATTTCGCTGCTGCTGAGCTGAAATCGGAT 480  
Qy 688 TCGTCCCGCTGTGCCAGGAGAGCAGCGTGGTCCGCACTGTGTGCTCTTCCCGCGGAC 747  
Db 481 TCGTCCCGCTGTGCCAGGAGAGCAGCGTGGTCCGCACTGTGTGCTCTTCCCGCGGAC 540  
Qy 748 CTGAGCTGCCGAGTCCGACGAGTGTGAGCTCTCCGGCTACGGCAAGCATGAGGCTTGG 807  
Db 541 CTGAGCTGCCGAGTCCGACGAGTGTGAGCTCTCCGGCTACGGCAAGCATGAGGCTTGG 600



```

5', mRNA sequence.
BQ278177
VERSION BQ278177.1 GI:20488385
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 993)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCW2043 row: 0 column: 10
High quality sequence stop: 681.
FEATURES
Location/Qualifiers
1..993
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5805153"
/tissue_type="teratocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC_109"
/notes="Organ: ovary; Vector: pOTB7; Site 1: EcoRI; Site 2:
XhoI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGGAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH MGC Library."
BASE COUNT 235 a 295 c 269 g 194 t
ORIGIN
Query Match 72.1%; Score 813.8; DB 13; Length 993;
Best Local Similarity 99.6%; Pred. No. 4.3e-201;
Matches 826; Conservative 0; Mismatches 2; Indels 1; Gaps 1;
QY 300 GACGTGGGAGTACTGTGATGGCCCTCTCTGCTCCACCTGGCGCTGACAGTACAGCCA 359
|
|
|
Db 1 GACGTGGGAGTACTGTGATGGCCCTCTCTGCTCCACCTGGCGCTGACAGTACAGCCA 59
|
|
|
QY 360 GCCTCAGTTTCGCATCAAGAGGGCTCTTCGCGACATCGCTCCACCCCTGGCAGGC 419
|
|
|
Db 60 GCCTCAGTTTCGCATCAAGAGGGCTCTTCGCGACATCGCTCCACCCCTGGCAGGC 119
|
|
|
QY 420 TGCCATCTTTGCAAGCAGAGAGTTCGCGGAGAGCGGTTCTCTGTCGGGGGCTACT 479
|
|
|
Db 120 TGCCATCTTTGCAAGCAGAGAGTTCGCGGAGAGCGGTTCTCTGTCGGGGGCTACT 179
|
|
|
QY 480 CATGAGTCTCTGATTTCTCTGCGCGCCCATCTGCTTCAGGAGAGGTTTCGCGCCCA 539
|
|
|
Db 180 CATGAGTCTCTGATTTCTCTGCGCGCCCATCTGCTTCAGGAGAGGTTTCGCGCCCA 239
|
|
|
QY 540 CCACCTGACGGTATCTTGGGAGAGATCATCGGGTGTCTCTGCGGAGGAGGAGGAGAA 599
|
|
|
Db 240 CCACCTGACGGTATCTTGGGAGAGATCATCGGGTGTCTCTGCGGAGGAGGAGGAGAA 299
|
|
|
QY 600 ATTTGAAGTCGAAAATACATTTGTCATAGAGATTCGATGATGACATTTACGACAATGA 659
|
|
|
Db 300 ATTTGAAGTCGAAAATACATTTGTCATAGAGATTCGATGATGACATTTACGACAATGA 359
|
|
|
QY 660 CATTGCGCTGCTGACGTGAATCGGATTCGTCGCTGTGCCAGGAGACGCGTGT 719
|
|
|

```

**ACCESSION** BQ278177  
**VERSION** BQ278177.1 GI:20488385  
**KEYWORDS** EST.  
**SOURCE** Homo sapiens (human)  
**ORGANISM** Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
**REFERENCE** 1 (bases 1 to 993)  
**AUTHORS** NIH-MGC http://mgi.nci.nih.gov/.  
**TITLE** National Institutes of Health, Mammalian Gene Collection (MGC)  
**JOURNAL** Unpublished  
**COMMENT** Contact: Robert Strausberg, Ph.D.  
 Email: cgapbs-r@mail.nih.gov  
 Tissue Procurement: ATCC  
 cDNA Library Preparation: Rubin Laboratory  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: LLCW2043 row: 0 column: 10  
 High quality sequence stop: 681.

**FEATURES**  
 Location/Qualifiers  
 1..993  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5805153"  
 /tissue\_type="teratocarcinoma, cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="NIH MGC\_109"  
 /note="Organ: ovary; Vector: pOTB7; Site 1: EcoRI; Site 2:  
 XhoI; cDNA made by oligo-dT priming. Directionally cloned  
 into EcoRI/XhoI sites using the following 5' adaptor:  
 GGCAGGAG(G). Library constructed by Ling Hong in the  
 laboratory of Gerald M. Rubin (University of California,  
 Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and  
 Superscript II RT (Life Technologies). Note: this is a  
 NIH MGC Library."

**BASE COUNT** 235 a 295 c 269 g 194 t  
**ORIGIN**

Query Match 72.1%; Score 813.8; DB 13; Length 993;  
 Best Local Similarity 99.6%; Pred. No. 4.3e-201;  
 Matches 826; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 300 GACGTGGGAGTACTGTGATGGCCCTCTCTGCTCCACCTGGCGCTGACAGTACAGCCA 359  
 |  
 |  
 |  
 Db 1 GACGTGGGAGTACTGTGATGGCCCTCTCTGCTCCACCTGGCGCTGACAGTACAGCCA 59  
 |  
 |  
 |  
 QY 360 GCCTCAGTTTCGCATCAAGAGGGCTCTTCGCGACATCGCTCCACCCCTGGCAGGC 419  
 |  
 |  
 |  
 Db 60 GCCTCAGTTTCGCATCAAGAGGGCTCTTCGCGACATCGCTCCACCCCTGGCAGGC 119  
 |  
 |  
 |  
 QY 420 TGCCATCTTTGCAAGCAGAGAGTTCGCGGAGAGCGGTTCTCTGTCGGGGGCTACT 479  
 |  
 |  
 |  
 Db 120 TGCCATCTTTGCAAGCAGAGAGTTCGCGGAGAGCGGTTCTCTGTCGGGGGCTACT 179  
 |  
 |  
 |  
 QY 480 CATGAGTCTCTGATTTCTCTGCGCGCCCATCTGCTTCAGGAGAGGTTTCGCGCCCA 539  
 |  
 |  
 |  
 Db 180 CATGAGTCTCTGATTTCTCTGCGCGCCCATCTGCTTCAGGAGAGGTTTCGCGCCCA 239  
 |  
 |  
 |  
 QY 540 CCACCTGACGGTATCTTGGGAGAGATCATCGGGTGTCTCTGCGGAGGAGGAGGAGAA 599  
 |  
 |  
 |  
 Db 240 CCACCTGACGGTATCTTGGGAGAGATCATCGGGTGTCTCTGCGGAGGAGGAGGAGAA 299  
 |  
 |  
 |  
 QY 600 ATTTGAAGTCGAAAATACATTTGTCATAGAGATTCGATGATGACATTTACGACAATGA 659  
 |  
 |  
 |  
 Db 300 ATTTGAAGTCGAAAATACATTTGTCATAGAGATTCGATGATGACATTTACGACAATGA 359  
 |  
 |  
 |  
 QY 660 CATTGCGCTGCTGACGTGAATCGGATTCGTCGCTGTGCCAGGAGACGCGTGT 719  
 |  
 |  
 |

**Db** 360 CATTGCGCTGCTGACGTGAATCGGATTCGTCGCTGTGCCAGGAGACGCGTGT 419  
**QY** 720 CCGCACTGTGTGCTTCCCGCGGAGACCTGCGACCTGCGGACCTGCGAGTGTGAGCT 779  
**Db** 420 CCGCACTGTGTGCTTCCCGCGGAGACCTGCGACCTGCGGAGTGTGAGCT 479  
**QY** 780 CTCGGCTAGCGCAAGCATGAGGCTTGTCTCTTTCTATTTCGAGGCGCTGAAGGAGGC 839  
**Db** 480 CTCGGCTAGCGCAAGCATGAGGCTTGTCTCTTTCTATTTCGAGGCGCTGAAGGAGGC 539  
**QY** 840 TCATGTGAGACTGTACCCATCCAGCCGCTGCACATCAACATTTACTTTACAGACAGT 899  
**Db** 540 TCATGTGAGACTGTACCCATCCAGCCGCTGCACATCAACATTTACTTTACAGACAGT 599  
**QY** 900 CACGACAAACATGTGTGTGTGTGAGACACTCGGAGCGGGGCCCGGCAAACTTTGCA 959  
**Db** 600 CACGACAAACATGTGTGTGTGTGAGACACTCGGAGCGGGGCCCGGCAAACTTTGCA 659  
**QY** 960 CGAGCCTGCGAGGCGGATTCGGAGGCGCCCTGCTGTGTCTGAACGATGCCGATGAC 1019  
**Db** 660 CGAGCCTGCGAGGCGGATTCGGAGGCGCCCTGCTGTGTCTGAACGATGCCGATGAC 719  
**QY** 1020 TTTGTTGGGCATCATCAGCTGGGGCTGGGCTGTGGACAGAGGATGTCCCGGCTGTGA 1079  
**Db** 720 TTTGTTGGGCATCATCAGCTGGGGCTGGGCTGTGGACAGAGGATGTCCCGGCTGTGA 779  
**QY** 1080 CACAAAGTTTACCAATACCTAGACTGGATTTCGTGACAAACATGCGACCG 1128  
**Db** 780 CACAAAGTTTACCAATACCTAGACTGGATTTCGTGACAAACATGCGACCG 828

**RESULT 6**  
**BO690978** 883 bp mRNA linear EST 15-JUL-2002  
**LOCUS** AGENCOURT 8046707 NIH\_MGC\_110 Homo sapiens cDNA clone IMAGE:6209120  
**DEFINITION** 5', mRNA sequence.  
**ACCESSION** BO690978  
**VERSION** BO690978.1 GI:21816294  
**KEYWORDS** EST.  
**SOURCE** Homo sapiens (human)  
**ORGANISM** Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
**REFERENCE** 1 (bases 1 to 883)  
**AUTHORS** NIH-MGC http://mgi.nci.nih.gov/.  
**TITLE** National Institutes of Health, Mammalian Gene Collection (MGC)  
**JOURNAL** Unpublished  
**COMMENT** Contact: Robert Strausberg, Ph.D.  
 Email: cgapbs-r@mail.nih.gov  
 Tissue Procurement: ATCC  
 cDNA Library Preparation: Rubin Laboratory  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: LLCW2367 row: 0 column: 09  
 High quality sequence stop: 558.

**FEATURES**  
 Location/Qualifiers  
 1..883  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:6209120"  
 /tissue\_type="ductal carcinoma, cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="NIH MGC\_110"  
 /note="Organ: pancreas; Vector: pOTB7; Site 1: XhoI;  
 Site 2: EcoRI; cDNA made by oligo-dT priming.  
 Directionally cloned into EcoRI/XhoI sites using the  
 following 5' adaptor: GGCAGGAG(G). Library constructed by  
 Ling Hong in the laboratory of Gerald M. Rubin (University  
 of California, Berkeley) using ZAP-cDNA synthesis kit

(Stratagene) and Superscript II RT (Life Technologies).  
Note: this is a NIH\_MGC Library."

```
BASE COUNT      184 a      265 c      259 g      175 t
ORIGIN
Query Match      71.9%; Score 811.2; DB 13; Length 883;
Best Local Similarity 99.0%; Pred. No. 1.9e-200;
Matches 816; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 178 GCACAGGTTTACACAGCAGACAGACCCAGTGCACAGGCACTGGGCTGGGCAACATAAT 237
Db 1 GGCAGAGTTTACACAGCAGACAGACCCAGTGCACAGGCACTGGGCTGGGCAACATAAT 60
QY 238 TACTGCCGGAATCCTGATGGGATCCCAAGCCCTGGTCCACAGTCTCTCAAGAACCGCAGG 297
Db 61 TACTGCCGGAATCCTGATGGGATCCCAAGCCCTGGTCCACAGTCTCTCAAGAACCGCAGG 120
QY 298 CTGACGTGGGAGTACTGTGATGTGCCCTCTCTGCTCCACTGCGGCTGAGACAGTACAGC 357
Db 121 CTGACGTGGGAGTACTGTGATGTGCCCTCTCTGCTCCACTGCGGCTGAGACAGTACAGC 180
QY 358 CAGCCTCAGTTTCGATCAAAAGGAGGCTCTTCCGCGACATCGCTCCACCCCTGGCAG 417
Db 181 CAGCCTCAGTTTCGATCAAAAGGAGGCTCTTCCGCGACATCGCTCCACCCCTGGCAG 240
QY 418 GCTGCCATCTTTGCCAAGCAGAGAGTCCGCCGAGAGCGGTTCTCTGCGGGGGCATA 477
Db 241 GCTGCCATCTTTGCCAAGCAGAGAGTCCGCCGAGAGCGGTTCTCTGCGGGGGCATA 300
QY 478 CTATCAGCTCTGTGTGATTTCTCTCTGCGGCCACTGCTTCCAGAGAGGTTTCCGCC 537
Db 301 CTATCAGCTCTGTGTGATTTCTCTCTGCGGCCACTGCTTCCAGAGAGGTTTCCGCC 360
QY 538 CACCACTCAGCTGATCTTGGCAGACATACCGGGTGTCCCTGGCGAGGAGGAGCAG 597
Db 361 CACCACTCAGCTGATCTTGGCAGACATACCGGGTGTCCCTGGCGAGGAGGAGCAG 420
QY 598 AAATTTGAAGTCGAAATAATATTCTCCATAGGAATTCGATGATGACATTTACACAAT 657
Db 421 AAATTTGAAGTCGAAATAATATTCTCCATAGGAATTCGATGATGACATTTACACAAT 480
QY 658 GACATTCGCTGCTCAGCTGAATTCGATTCGTCCTCCGCTGCTCCAGAGAGAGCGTG 717
Db 481 GACATTCGCTGCTCAGCTGAATTCGATTCGTCCTCCGCTGCTCCAGAGAGAGCGTG 540
QY 718 CTCGCCACTGTGTGCTTTCCCGCGGACCTGACGCTCCGAGTGAACGAGTGTGAG 777
Db 541 CTCGCCACTGTGTGCTTTCCCGCGGACCTGACGCTCCGAGTGAACGAGTGTGAG 600
QY 778 CTCCTCCGCTACGGCAAGCATGAGGCTTTGCTCTCTTTCTATTTCGAGCGGCTGAAGGAG 837
Db 601 CTCCTCCGCTACGGCAAGCATGAGGCTTTGCTCTCTTTCTATTTCGAGCGGCTGAAGGAG 660
QY 838 GCTCATGTGACAGTGTACCCATCCAGCCGCTGCAATCACAACTTTACTTAAACAGACA 897
Db 661 GCTCATGTGACAGTGTACCCATCCAGCCGCTGCAATCACAACTTTACTTAAACAGACA 720
QY 898 GTCACCGACACATGCTGTGCTGGAGACACTCGAGGGGGGGGGCCCGACCAACTTG 957
Db 721 GTCACCGACACATGCTGTGCTGGAGACACTCGAGGGGGGGGGCCCGACCAACTTG 780
QY 958 CACGACGCTGCGACGGGATTCGGAGGGCCCTCTGGTGTCT 1001
Db 781 CACGACGCTGCGACGGGATTCGGAGGGCCCTCTGGTGTGT 824
```

RESULT 7  
BX389608  
LOCUS BX389608 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA  
DEFINITION clone CS0D1039YK15 5-PRIME, mRNA sequence.  
ACCESSION BX389608  
VERSION BX389608.1 GI:30463469

KEYWORDS  
SOURCE  
ORGANISM

EST.  
Homo sapiens (human)  
Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1. (bases 1 to 918)  
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.  
Full-length cDNA libraries and normalization  
Unpublished

Contact: Genoscope

Genoscope - Centre National de Sequencage

BP 191 91006 EVRY cedex - France

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr

was normalized. Library was constructed by Life Technologies, a  
division of Invitrogen. This sequence belongs to sequence cluster  
4498.r For more information about this cluster, see

http://www.genoscope.cns.fr/

cgi-bin/cluster.cgi?seq=CS0BAP025ZG09\_AF02376\_2&cluster=4498.r.

Contact : Feng Liang Email : fliang@lifetech.com URL :

http://fulllength.invitrogen.com/Invitrogen Corporation 1600

Paraday Avenue Genoscope sequence ID : CS0BAP025ZG09\_AF02376\_2.

Location/Qualifiers

source

1. 918

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="CS0D1039YK15"

/tissue\_type="PLACENTA COT 25-NORMALIZED"

/clone\_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"

/notes="1st strand cDNA was primed with a NotI-oligo (dT)

primer. Five prime end enriched, double-strand cDNA was

digested with Not I and cloned into the Not I and EcoR V

sites of the pCMVSPORT 6 vector. Library was normalized."

BASE COUNT 189 a 273 c 264 g 186 t 6 others

ORIGIN

Query Match

Best Local Similarity 96.4%; Pred. No. 1.7e-199;

Matches 855; Conservative 0; Mismatches 29; Indels 3; Gaps 3;

QY 236 ATTACTGCCGGAATCTTGATGGGATGCCAAGCCCTGGTCCAGCTGCTGAAGAACCGCA 295

Db 17 ATTACTGCCGGAATCTTGATGGGATGCCAAGCCCTGGTCCAGCTGCTGAAGAACCGCA 76

QY 296 GGCTGACGTGGGAGTACTGTGATGTGCCCTCTCTCCACTCGCGCTCAGACAGTACA 355

Db 77 GGCTGACGTGGGAGTACTGTGATGTGCCCTCTCTCCACTCGCGCTCAGACAGTACA 136

QY 356 GCCAGCCTCAGTTTTCGATCAAAAGGAGGCTCTTTCGCCGACATCGCTCCACCCCTGGC 415

Db 137 GCCAGCCTCAGTTTTCGATCAAAAGGAGGCTCTTTCGCCGACATCGCTCCACCCCTGGC 196

QY 416 AGGCTGCATCTTTGCCAAGCAGAGAGTCCGCCGAGAGCGGTTCTGTGCGGGGGCA 475

Db 197 AGGCTGCATCTTTGCCAAGCAGAGAGTCCGCCGAGAGCGGTTCTGTGCGGGGGCA 256

QY 476 TACTCATCAGCTCTGTGATTTCTCTCGCCGCCACTGCTTCCAGAGAGGTTTCCGC 535

Db 257 TACTCATCAGCTCTGTGATTTCTCTCGCCGCCACTGCTTCCAGAGAGGTTTCCGC 316

QY 536 CCCACCACTTCAGCGGTGATCTTGGCGAGAACATACCGGGTGGTCCCTGGCGAGGAGGC 595

Db 317 CCCACCACTTCAGCGGTGATCTTGGCGAGAACATACCGGGTGGTCCCTGGCGAGGAGGC 376

QY 596 AGAAATTTGAAGTCGAAAAATACATTGTCCATAGGAATTCGATGATGACACTTACACA 655

Db 377 AGAAATTTGAAGTCGAAAAATACATTGTCCATAGGAATTCGATGATGACACTTACACA 436

QY 656 ATGACATTCGCTGCTGACGTGAAATTCGATTCGTCCTCCCTGTGCCAGAGAGCAGCG 715

Db 437 ATGACATTCGCTGCTGACGTGAAATTCGATTCGTCCTCCCTGTGCCAGAGAGCAGCG 496

QY 716 TGGTCCGCACTGTGTGCTTTCCCGCGGAGACCTGCAGCTGCCGAGGTGTGTG 775

```
Db 497 TGGTCCGACATGTGTGCTTCCCCCGCGGACCTTCAGCTGCCGACTGGACGGAGTGTG 556
Qy 776 AGCTCTCCGGCTACCGCAAGCATGAGGCTTGTCTCTCTTCTATTTCGAGCGGCTGAAGG 835
Db 557 AGCTCTCCGGCTACCGCAAGCATGAGGCTTGTCTCTCTTCTATTTCGAGCGGCTGAAGG 616
Qy 836 AGGCTCATGTGAGTGTACCATGACCGCTGACATCAACAATTTACTTAACAGAA 895
Db 617 AGGCTCATGTGAGTGTACCATGACCGCTGACATCAACAATTTACTTAACANGA 676
Qy 896 CAGTCACCGACAAATGTGTGTGCTGGAGACACTCGGAGCGGGCCCGAGCAAACT 955
Db 677 CAGTCACCGACAAATGTGTGTGCTGGAGACACTCGNAGCGGGCCGCCCA-GCAAACT 735
Qy 956 TGCACGACGCTCCGACGCGGATTCGGGAGGCCCCCTGCTGTGCTGAACATGGCCGCA 1015
Db 736 TGCACGACGCTCCGACGCGGATTCGCGAGGCCCCCTGCTGTGCTTTCAGCATGGCCAA 795
Qy 1016 TGA-CTTTGGTGGGATCATGCTGAGTGGG-CTGGGCTGTGGACAGAAAGATGCCGG 1073
Db 796 TGACCTTTGGTGGGATCATGCTGAGTGGGCTGGGCTGTTCACAGAAAGGATGTTCCGG 855
Qy 1074 TGTGTACAAAGGTTTACCACTACCTAGATGGATTCGTGACACA 1120
Db 856 TGGTGTCAACCAAGTTTAACTACCTAAACCGGATTCGGGACACA 902
```

```
RESULT 8
BU845014
LOCUS
DEFINITION BU845014 930 bp mRNA linear EST 16-OCT-2002
IMAGE:6578533 5', mRNA sequence.
ACCESSION BU845014.1 GI:24029455
VERSION EST.
KEYWORDS Homo sapiens (human)
SOURCE
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS 1 (bases 1 to 930)
TITLE NIH-MGC http://mgi.nci.nih.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-re@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
Clone distribution by: Agencourt Bioscience Corporation
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM2780 row: 9 column: 13
High quality sequence stop: 639.
Location/Qualifiers
1. 930
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6578533"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_109"
/note="Organ: ovary; Vector: pOTB7; Site 1: EcoRI; Site 2:
XhoI; cDNA made by oligo-dt priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCG(G). Library constructed by Ling Hong in the
Laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene), and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC Library."
```

BASE COUNT 199 a 279 c 266 g 186 t

## ORIGIN

```
Query Match 70.6%; Score 796.2; DB 13; Length 930;
Best Local Similarity 99.4%; Pred. No. 1.6e-196;
Matches 820; Conservative 0; Mismatches 3; Indels 2; Gaps 2;
Qy 300 GACGTGGAGTACTGTGATGTGCTTCCCTGCTCCACCTGCGGCTGAGACAGTACAGCA 359
Db 1 GACGTGGAGTACTGTGATGTGCTTCCCTGCTCCACCTGCGGCTGAGACAGTACAGCA 60
Qy 360 GCCTTCAGTTTCGCATCAAGAGGAGGCTTTCGCCGACATCGCTCCACCCCTCGCAGGC 419
Db 61 GCCTTCAGTTTCGCATCAAGAGGAGGCTTTCGCCGACATCGCTCCACCCCTCGCAGGC 120
Qy 420 TGCCATCTTTGCCAAGCAGAGGAGTCCGCCGAGAGCGGTTCCTGTGCGGGGCACTACT 479
Db 121 TGCCATCTTTGCCAAGCAGAGGAGTCCGCCGAGAGCGGTTCCTGTGCGGGGCACTACT 180
Qy 480 CATCAGCTCTGCTGGATTCCTCTGCGGCGCACTGCTTCAGGAGAGGTTTCGCGCCCA 539
Db 181 CATCAGCTCTGCTGGATTCCTCTGCGGCGCACTGCTTCAGGAGAGGTTTCGCGCCCA 240
Qy 540 CCACCTCAGCGTGTATCTTGGGCGAGAAATACCGGGTGTCCCTGCGGAGGAGGAGCA 599
Db 241 CCACCTCAGCGTGTATCTTGGGCGAGAAATACCGGGTGTCCCTGCGGAGGAGGAGCA 300
Qy 600 ATTTGAAGTGCAGAAATACATTTGTCCATAAGGAATTCGATGATGACATTCACGACAATGA 659
Db 301 ATTTGAAGTGCAGAAATACATTTGTCCATAAGGAATTCGATGATGACATTCACGACAATGA 360
Qy 660 CATTGCGCTGCTGAGCTGAAATCGAATTCGTCCTGCTGTCGCCAGAGAGCAGCGTGGT 719
Db 361 CATTGCGCTGCTGAGCTGAAATCGAATTCGTCCTGCTGTCGCCAGAGAGCAGCGTGGT 420
Qy 720 CCGCAGCTGTGCTTCCCGCGGAGCCTGCGAGTCCCGAGTGGAGCGGAGTGTGAGCT 779
Db 421 CCGCAGCTGTGCTTCCCGCGGAGCCTGCGAGTCCCGAGTGGAGCGGAGTGTGAGCT 480
Qy 780 CTCGCGCTACGGCAAGCATGAGGCTTGTCTCTCTCTTCTATTTCGAGCGGCTGGAAGGAGGC 839
Db 481 CTCGCGCTACGGCAAGCATGAGGCTTGTCTCTCTTCTATTTCGAGCGGCTGGAAGGAGGC 540
Qy 840 TCATGTGAGTGTGATCCCATCCAGCGCTGACATCAACAATTTACTTTAAAGAAACAGT 899
Db 541 TCATGTGAGTGTGATCCCATCCAGCGCTGACATCAACAATTTACTTTAAAGAAACAGT 600
Qy 900 CACCGCAACATGCTGTGCTGGAGACATTCGGAGCGGGCGGCCCGGCAAACTTGCA 959
Db 601 CACCGCAACATGCTGTGCTGGAGACATTCGGAGCGGGCGGCCCGGCAAACTTGCA 660
Qy 960 CGAGCGCTGCGAGGCGGATTCGGAGGCGGCCCTGCTGTCTGAAACGATGCGCCGATGAC 1019
Db 661 CGAGCGCTGCGAGGCGGATTCGGAGGCGGCCCTGCTGTCTGAAACGATGCGCCGATGAC 720
Qy 1020 TTTGGTGGGATCATCAGCTGGGCGCTGGGCTGTGGAGAGAGGATTCCTC-GGGTGTGT 1078
Db 721 TTTGGTGGGATCATCAGCTGGGCGCTGGGCTGTGGAGAGAGGATTCCTC-GGGGCTGTGT 780
Qy 1079 ACA-CAAGGTTACCAACTACCTAGACTGGATTCGTGACACATG 1122
Db 781 ACACCAAGGGTACCAACTACCTAGACTGGATTCGTGACACATG 825
```

## RESULT 9

```
BU157720
LOCUS
DEFINITION BU157720 888 bp mRNA linear EST 04-SEP-2002
IMAGE:6578533 5', mRNA sequence.
ACCESSION BU157720
VERSION BU157720.1 GI:22671252
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
```









QY 772 TGTGAGCTCTCCGGCTACGCAAGCATGAGGCTTGTCTCTCTTTCTATTTCGGAGCGGCTG 831  
Db 421 TGTGAGCTCTCCGGCTACGCAAGCATGAGGCTTGTCTCTCTTTCTATTTCGGAGCGGCTG 480  
QY 832 AAGGAGGCTCATGTGAGCTGTACCCATCCAGCCGCTGCACATCAACAATTTACTTTAAC 891  
Db 481 AAGGAGGCTCATGTGAGCTGTACCCATCCAGCCGCTGCACATCAACAATTTACTTTAAC 540  
QY 892 AGAAGAGTCACGCAACATGCTGTGCTGTGAGACATCCGAGCGGGGGCCCCAGGCA 951  
Db 541 AGAAGAGTCACGCAACATGCTGTGCTGTGAGACATCCGAGCGGGGGCCCCAGGCA 600  
QY 952 AACTTGCACGAGCCCTGCGAGGGGCAATTCGGAGGCCCCCTGCTGTCTGAAACATGGC 1011  
Db 601 AACTTGCACGAGCCCTGCGAGGGGCAATTCGGAGGCCCCCTGCTGTCTGAAACATGGC 660  
QY 1012 GCATGACTTTGGTGGGCAATCATACGTGGGCGCTGGGCTGTGGGACAGAGATGTCCCG 1071  
Db 661 GCATGACTTTGGTGGGCAATCATACGTGGGCGCTGGGCTGTGGGACAGAGATGTCCCG 720  
QY 1072 GGTGTGTACACAAAGGTTACCAACTACCTAGAC-TGGATTGCTGACACATGCCGACCG 1128  
Db 721 GGTGTGTACACAAAGGTTACCAACTACCTAGACTTGGATTGCTGACACATGCCGACCG 778

RESULT 12  
LOCUS BE616613  
DEFINITION 601278964F1 NIH\_MGC\_39 Homo sapiens cDNA clone IMAGE:3611007 5',  
mRNA sequence.  
ACCESSION BE616613  
VERSION BE616613.1 GI:9898212  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE 1 (bases 1 to 948)  
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-re@mail.nih.gov](mailto:cgapbs-re@mail.nih.gov)  
Tissue procurement: ATCC  
cDNA Library Preparation: Ling Hong/Rubin Laboratory  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: L1CM267 row: P column: 16  
High quality sequence start: 2  
High quality sequence stop: 759  
Location/Qualifiers  
1. 948  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:3611007"  
/tissue\_type="adenocarcinoma"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH MGC 39"  
/note="Organ: pancreas; Vector: pOTB7; Site\_1: XhoI; Site\_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."

BASE COUNT 173 a 295 c 293 g 187 t  
ORIGIN

Query Match 62.7%; Score 706.8; DB 10; Length 948;  
Best Local Similarity 93.3%; Pred. No. 3.2e-173;  
Matches 793; Conservative 0; Mismatches 52; Indels 5; Gaps 5;  
QY 72 GGGAAACAGTGACTGTCTACT-TTGGGAATGGGTGAGCTACCGTGGGACGACAGCCTCA 130  
Db 8 GAGGAAACAGTGACTGTCTATTATTGGGAATGGGTGAGCTACCGTGGGACGACAGCCTCA 67  
QY 131 CCGAGTGGGTGGCTCTCTGCTCCG-CTGGNAATTCATGATCTCTGATAGGCAAGGTTTAC 189  
Db 68 CCGAGTGGGTGGCTCTCTGCTCCGCTCCGCTGGGAATTCATGATCTCTGATAGGCAAGGTTTAC 127  
QY 190 ACAGCACAAGAACCCAGTGGCCAGGCACTGGGCTGGGCAAAACATAATTAATCTGCGGAAT 249  
Db 128 ACAGCACAAGAACCCAGTGGCCAGGCACTGGGCTGGGCAAAACATAATTAATCTGCGGAAT 187  
QY 250 CCTGATGGGATGCCAAGCCCTGGTGGCAAGCTGTGAAGAACCCAGGCTGACGTGGGAG 309  
Db 188 CCTGATGGGATGCCAAGCCCTGGTGGCAAGCTGTGAAGAACCCAGGCTGACGTGGGAG 247  
QY 310 TACTGTGATGTGCTCTCTGCTCCACCTGGGCTGGAGACAGTACAGCCAGCCTCAGTTT 369  
Db 248 TACTGTGATGTGCTCTCTGCTCCACCTGGGCTGGAGACAGTACAGCCAGCCTCAGTTT 307  
QY 370 CGCATCAAGAGGAGGCTCTTTCGCCGACATCGCCCTCCACCCCTGGCAGGCTGCCATCTTT 429  
Db 308 CGCATCAAGAGGAGGCTCTTTCGCCGACATCGCCCTCCACCCCTGGCAGGCTGCCATCTTT 367  
QY 430 GCCAAGCACAAGGAGGTGCGCCGGAGAGCGGTTCCTGTGCGGGGGCATACTCATCAGCTCC 489  
Db 368 GCCAAGCACAAGGAGGTGCGCCGGAGAGCGGTTCCTGTGCGGGGGCATACTCATCAGCTCC 427  
QY 490 TGTGAGATTTCTCTGCGGGCCACTGCTTCAGAGAGAGGTTTCCGCCCCACCACCTGACG 549  
Db 428 TGTGAGATTTCTCTGCGGGCCACTGCTTCAGAGAGAGGTTTCCGCCCCACCACCTGACG 487  
QY 550 GTGATCTTGGGCAACATACCGGGTGGTCCCTGGCAGGAGGAGCAGAAATTTGAAAGTC 609  
Db 488 GTGATCTTGGGCAACATACCGGGTGGTCCCTGGCAGGAGGAGCAGAAATTTGAAAGTC 547  
QY 610 GAAAAATACATTTGCCATAAGGAATTCGATGATGACACTTACGACAAATGACATTTGCGCTG 669  
Db 548 GAAAAATACATTTGCCATAAGGAATTCGATGATGACACTTACGACAAATGACATTTGCGCTG 607  
QY 670 CTGAGCTGAAATCGGATTTGTCCTGCTGCGCCAGGAGAGCAGCGTGGTCCGCACTGTG 729  
Db 608 CTGAGCTGAAATCGGATTTGTCCTGCTGCGCCAGGAGAGCAGCGTGGTCCGCACTGTG 667  
QY 730 TGCCTTCCCGCGGACCTGCGAGCTGCGAGCTGGAGCGGAGTGTGAGCTCTCCGGCTAC 789  
Db 668 TGCCTT-CCCGCGGAGCTGCGAGCTGCGGAGCGGAGTGTGAGCTCTCCGGCTAC 726  
QY 790 GGCAAGCATGAGCTTTGTCTCTCTTTCTATTTCGGAGCGGCTGGAAGAGGCTCATGTCTAGA 849  
Db 727 GGCAAGCATGAGG-CTTGTCTCTCTTTCTATTTCGGAGCGGCTGGAAGAGGCTCATGTCTAGA 785  
QY 850 CTGTACCCATCCAGCCGCTGCACATCAACAATTTACTTAACAGAACAGTACCAGCAAC 909  
Db 786 CTG-GTCCATCCAGCCGCTGGGCTTCCGCCACCTTTCTTTCAGGACAGCGCGCGCTGC 844  
QY 910 ATGCTGTGTG 919  
Db 845 TGGGCTGGG 854

RESULT 13  
LOCUS BG763582  
DEFINITION 602735987F1 NIH\_MGC\_49 Homo sapiens cDNA clone IMAGE:4861268 5',  
mRNA sequence.  
ACCESSION BG763582  
VERSION BG763582.1 GI:14074235  
KEYWORDS EST.

SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 958)  
 AUTHORS NIH-MGC <http://mgi.nci.nih.gov/>  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: [cgabbs-r@mail.nih.gov](mailto:cgabbs-r@mail.nih.gov)  
 Tissue Procurement: ATCC/DCTD/DTF  
 cDNA Library Preparation: Ling Hong/Rubin Laboratory  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone Distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
 Plate: LLCM1719 row: n column: 21  
 High quality sequence stop: 810.

FEATURES  
 Location/Qualifiers  
 1..958  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:4861268"  
 /tissue\_type="melanotic melanoma, high MDR (cell line)"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="NIH MGC 49"  
 /note="Organ: skin; Vector: pOTB7; Site 1: XhoI; Site 2:  
 EcoRI; cDNA made by oligo-dr priming. Directionally cloned  
 into EcoRI/XhoI sites using the following 5' adaptor:  
 GGCCAGG(G). Size-selected >500bp for average insert size  
 1.8kb. Library constructed by Ling Hong in the laboratory  
 of Gerald M. Rubin (University of California, Berkeley)  
 using ZAP-cDNA synthesis kit (Stratagene) and Superscript  
 II RT (Life Technologies). Note: this is a NIH\_MGC  
 Library. |"  
 BASE COUNT 198 a 285 c 282 g 193 t  
 ORIGIN  
 Query Match 62.1%; Score 700.4; DB 12; Length 958;  
 Best Local Similarity 93.8%; Fred. No. 1.5e-171;  
 Matches 840; Conservative 0; Mismatches 41; Indels 15; Gaps 10;  
 QY 56 CCCAGCGGCTCTGAGGAAACAGTACTGCTACTTTGGGAATGGGTACGCTACCGCTG 115  
 DB 61 CCCCTGCTCTGAGGAAACAGTACTGCTACTTTGGGAATGGGTACGCTACCGCTG 120  
 QY 116 GCACGACAGCTTACACGAGTCGGTGCTCTGCTCCCGTGGAAATTCATGATCCTGA 175  
 DB 121 GCACGACAGCTTACACGAGTCGGTGCTCTGCTCCCGTGGAAATTCATGATCCTGA 180  
 QY 176 TAGGCAAGGTTTACACACAGACACGACCCAGTCCCGACGCTGGGCTGGGCAACATA 235  
 DB 181 TAGGCAAGGTTTACACACAGACACGACCCAGTCCCGACGCTGGGCTGGGCAACATA 240  
 QY 236 ATTACTGCGGAATCCTGATGGGATGCCAAGCCCTGGTGCCACGCTGCTGAAGAACCGCA 295  
 DB 241 ATTACTGCGGAATCCTGATGGGATGCCAAGCCCTGGTGCCACGCTGCTGAAGAACCGCA 300  
 QY 296 GGCTGACGTGGAGTACTGTGATGTCCTCTGCTCCACCTGGCGGCTGAGACAGTACA 355  
 DB 301 GGCTGACGTGGAGTACTGTGATGTCCTCTGCTCCACCTGGCGGCTGAGACAGTACA 360  
 QY 356 GCCAGCTCAGTTTCGATCAAGAGAGGGCTTTCGCCGACATGCTCCACCCCTGGC 415  
 DB 361 GCCAGCTCAGTTTCGATCAAGAGAGGGCTTTCGCCGACATGCTCCACCCCTGGC 420  
 QY 416 AGGCTGCCATCTTTGCCAAGCACAGAGGTGCGCCGGAGAGCGGTTCCTGTGCGGGGCA 475  
 DB 421 AGGCTGCCATCTTTGCCAAGCACAGAGGTGCGCCGGAGAGCGGTTCCTGTGCGGGGCA 480  
 QY 476 TACTCATCAGTCTCTGTGGATTCTCTCTGCGGCCCACTGCTTCCAGGAGAGGTTTCCGC 535

481 TACTCATCAGTCTCTGTGGATTCTCTCTGCGGCCCACTGCTTCCAGGAGAGGTTTCTCTG 540  
 QY 536 CCCACCACTGACCGTGTATCTTTGGGCAAGACATACCGGTGTCTCCCTGGCAGAGGAGC 595  
 DB 541 CCCACCACTGACCGTGTATCTTTGGGCAAGACATACCGGTGTCTCCCTGGCAGAGGAGC 600  
 QY 596 AGAAATTTGAAGTCGAAAAATACATTTGTCCATAAGGAATTCGATGATGACACTTACGACA 655  
 DB 601 AGAAATTTGAAGTCGAAAAATACATTTGTCCATAAGGAATTCGATGATGACACTTACGACA 660  
 QY 656 ATGACATTGGCTGCTGCTGACGCTGAATTCGATTTCGT-CCGCTGTGCCAGGAGAGCAGC 714  
 DB 661 ATGACATTGGCTGCTGCTGACGCTGAATTCGATTTCGT-CCGCTGTGCCAGGAGAGCAGC 720  
 QY 715 -GTGCTGCGCACTGTGTGCTTT--CCGCGCGGACCTGACGCTGCGGAGCTGCAC--GGA 770  
 DB 721 GTGTGCTGCACTGTGTGCTTTCCCTCCCGCGGACCTGACGCTGCGGAGCTGCACGGA 780  
 QY 771 GTGTGAGCTCTCCG-GCTACGGCAAGCAT--GAGGCGCTTGTCTCTTTCTATTTCGAGCG 827  
 DB 781 GTGTGAGCTCTCCGAGCTAGGCAAGCATTTGAGGCGCTGTCTCTCTTCTATTTCGAGCG 840  
 QY 828 GCTGAGGAGCTCATGTGACACTGTACCCATCCAGCGCTGACACATCACAACATTTACT 887  
 DB 841 G-TGAAGGAGGCTCATG--CGACTGGACCCATTCACCGGTGGCCATTTCACAACCTTTACT 897  
 QY 888 TAACAGAACAGTACCCGACACATGCTGTGTGTGGAGACACTCGGAGCGCGCGGC 943  
 DB 898 T-ACAGAACAGTACC---GAAATTTGTGTGTGCGGACCCCGACGCGGGGCCCCAGGC 949

RESULT 14  
 BO960216  
 LOCUS BO960216 974 bp mRNA linear EST 21-AUG-2002  
 DEFINITION AGENCOURT\_8945604 NIH\_MGC\_101 Homo sapiens cDNA clone IMAGE:6459629  
 5', mRNA sequence.  
 ACCESSION BO960216  
 VERSION BO960216.1 GI:22375694  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 974)  
 AUTHORS NIH-MGC <http://mgi.nci.nih.gov/>  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: [cgabbs-r@mail.nih.gov](mailto:cgabbs-r@mail.nih.gov)  
 Tissue Procurement: ATCC  
 cDNA Library Preparation: Rubin Laboratory  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone Distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
 Plate: LLCM645 row: a column: 06  
 High quality sequence stop: 603.

FEATURES  
 Location/Qualifiers  
 1..974  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:6459629"  
 /tissue\_type="epidermoid carcinoma, cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="NIH\_MGC\_101"  
 /note="Organ: lung; Vector: pOTB7; Site 1: EcoRI; Site 2:  
 XhoI; cDNA made by oligo-dr priming. Directionally cloned  
 into EcoRI/XhoI sites using the following 5' adaptor:  
 GGCCAGG(G). Library constructed by Ling Hong in the  
 laboratory of Gerald M. Rubin (University of California,

Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH MGC Library."

```
BASE COUNT      215 a 293 c 279 g 187 t
ORIGIN
Query Match      61.6%; Score 694.4; DB 13; Length 974;
Best Local Similarity 97.8%; Pred. No. 5.5e-170;
Matches 757; Conservative 0; Mismatches 11; Indels 6; Gaps 5;
QY 342 CTTGAGACAGTACAGCCAGCTCAGTTTCGATCAAAAGAGGGCTCTTCGCCGACATCGC 401
Db 19 CTTGAGACAGTACAGCCAGCTCAGTTTCGATCAAAAGAGGGCTCTTCGCCGACATCGC 78
QY 402 CTCACACCCCTGGCAGGCTGCATCTTTCGCAAGCAGAGGAGTCCGCCGAGAGCGGTT 461
Db 79 CTCACACCCCTGGCAGGCTGCATCTTTCGCAAGCAGAGGAGTCCGCCGAGAGCGGTT 138
QY 462 CTTGCGGGGGCATACATCAGCTCTCTGTGGATTCTCTGCGGCCCATCTGTTTCCA 521
Db 139 CTTGCGGGGGCATACATCAGCTCTCTGTGGATTCTCTGCGGCTC-CTGCTTCCA 197
QY 522 GGAGAGGTTTCGCCGCCACCACTGACCGTGAATCTTGGGCGAACAATACCGGGTGTGCC 581
Db 198 GGAGAGGTTTCGCCGCCACCACTGACCGTGAATCTTGGGCGAACAATACCGGGTGTGCC 257
QY 582 TGGCAGAGGAGGAGCAAAATTTGAAGTCGAAATACATTTGCCATAGGAATTCGATGA 641
Db 258 TGGCAGAGGAGGAGCAAAATTTGAAGTCGAAATACATTTGCCATAGGAATTCGATGA 317
QY 642 TGACACTTTACCAATGATTCGCTGCTGAGCTGAAATCGAATTCGTCCTCCGCTGTC 701
Db 318 TGACACTTTACCAATGATTCGCTGCTGAGCTGAAATCGAATTCGTCCTCCGCTGTC 377
QY 702 CCAGAGAGCAGCGTGTGTCGACATGTGTGCTTCCCGCGGAGCACTGCAGCTGCCGGA 761
Db 378 CCAGAGAGCAGCGTGTGTCGACATGTGTGCTTCCCGCGGAGCACTGCAGCTGCCGGA 437
QY 762 CTGAGCAGAGTGTAGCTCTCGGCTACGCAAGCATCAGAGCTGTCTCTCTTTCTATTC 821
Db 438 CTGAGCAGAGTGTAGCTCTCGGCTACGCAAGCATCAGAGCTGTCTCTCTTTCTATTC 497
QY 822 GGAGCGGCTGAAGAGGCTCATGTGAGACTGTACCCATCCAGCCGCTGCACATCACACA 881
Db 498 GGAGCGGCTGAAGAGGCTCATGTGAGACTGTACCCATCCAGCCGCTGCACATCACACA 557
QY 882 TTTACTTAACAGAACAGTACCGAACAATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 941
Db 558 TTTACTTAACAGAACAGTACCGAACAATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 617
QY 942 GCGCCAGGCAAACTGACAGCGCTGCGGCTGCGGAGGATTCGGAGG-CCCCCTGTGTGTC 1000
Db 618 GCGCCAGGCAAACTGACAGCGCTGCGGCTGCGGAGGATTCGGAGGCCCCCTGTGTGTC 677
QY 1001 TGAACGATGGCGCAT-GACTTTTGTGGGCTATCATCAGCT-GGGGCGCTGGGCTGTGGACA 1058
Db 678 TGAACGATGGCGCATGACTTTTGTGGGCTATCATCAGCTGGGCGCTGTGGACA 737
QY 1059 GAAGATGTCCC--GGGTGTGTACAAAGGTTTACCACTACCTAGACTGAATT 1110
Db 738 GAAGATGTCCC--GGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 791
```

```
RESULT 15
BE732704
LOCUS BE732704 709 bp mRNA linear EST 15-SEP-2000
DEFINITION 601571366F1 NIH_MGC_21 Homo sapiens cDNA clone IMAGE:392560 5',
mRNA sequence.
ACCESSION BE732704
VERSION BE732704.1 GI:10146696
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 709)  
NIH-MGC <http://mgc.nci.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished  
Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)  
Tissue Procurement: ATCC  
cDNA Library Preparation: Ling Hong/Rubin Laboratory  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: [image.llnl.gov](http://image.llnl.gov)  
Plate: L1CM752 row: d column: 17  
High quality sequence start: 3  
High quality sequence stop: 701.

#### FEATURES

source

1..709  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:392560"  
/tissue\_type="choriocarcinoma"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH\_MGC\_21"  
/note="Organ: placenta; Vector: pOTB7; Site\_1: XhoI; Site\_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."

BASE COUNT 147 a 219 c 203 g 140 t

ORIGIN

Query Match 61.5%; Score 693.4; DB 10; Length 709;  
Best Local Similarity 99.7%; Pred. No. 8.9e-170;  
Matches 705; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 263 CCAAGCCCTGGTGCACGCTGTGAAGAACCCGAGGCTGACGTGGAGTACTGTGATGTC 322  
Db 1 CCAAGCCCTGGTGCACGCTGTGAAGAACCCGAGGCTGACGTGGAGTACTGTGATGTC 60  
QY 323 CTTCTGCTCGACTGCGGCTGAGACAGTACAGCAGCCTCAGTTTCGCATCAAAGGAG 382  
Db 61 CTTCTGCTCGACTGCGGCTGAGACAGTACAGCAGCCTCAGTTTCGCATCAAAGGAG 120  
QY 383 GGCTTTTCGCGCATCGCTCCACCCCTGGCAGGCTGCCATCTTTGCCAAGCACAGGA 442  
Db 121 GGCTTTTCGCGCATCGCTCCACCCCTGGCAGGCTGCCATCTTTGCCAAGCACAGGA 180  
QY 443 GGTGCGCCGAGAGCGGTTCTCTGTGCGGGGCGATCATCATCAGCTCTCTGCTGGATTCTCT 502  
Db 181 GGTGCGCCGAGAGCGGTTCTCTGTGCGGGGCGATCATCATCAGCTCTCTGCTGGATTCTCT 240  
QY 503 CTGCGGCCCACTGTTCCAGAGAGGTTTCGCCGCCACCACTGACCGGTGATCTTGGGCA 562  
Db 241 CTGCGGCCCACTGTTCCAGAGAGGTTTCGCCGCCACCACTGACCGGTGATCTTGGGCA 300  
QY 563 GAACATACCGGGTGTCCCTGGGAGGAGGAGCAAGAAATTTGAAGTGGAAATATACATTG 622  
Db 301 GAACATACCGGGTGTCCCTGGGAGGAGGAGCAAGAAATTTGAAGTGGAAATATACATTG 360  
QY 623 TCCATAAGGAATTCGATGATGACATTTACGACAATGACATTCGCGCTGCTGAGCTGAAT 682  
Db 361 TCCATAAGGAATTCGATGATGACATTTACGACAATGACATTCGCGCTGCTGAGCTGAAT 420  
QY 683 CGGATTCGTCCTCCGCTGTGCCAGAGAGCAGCGTGTGCCGACTGTGTGCTTCCCGGG 742  
Db 421 CGGATTCGTCCTCCGCTGTGCCAGAGAGCAGCGTGTGCCGACTGTGTGCTTCCCGGG 480

Qy 743 CGGACCTGCAGCTGCCGACTGGACGGAGTGTGAGCTCTCCGGCTACCGCAAGCATGAGG 802  
Db |||||  
Qy 481 CGGACCTGCAGCTGCCGACTGGACGGAGTGTGAGCTCTCCGGCTACCGCAAGCATGAGG 540  
Db |||||  
Qy 803 CCTTGTCTCCTTTCTATTCCGGAGCGGCTGAAGGAGGCTCATGTCAAGACTGTACCCATCCA 862  
Db |||||  
Qy 541 CCTTGTCCCTTTCTATTCCGGAGCGGCTGAAGGAGGCTCATGTCAAGACTGTACCCATCCA 600  
Db |||||  
Qy 863 GCCGCTGCACATCACAACTTTACTTTAAACAGAAACAGTCACCGACAACATGCTGTGTGCTG 922  
Db |||||  
Qy 601 GCCGCTGCACATCACAACTTTACTTTACAGAACAGTCACCGACAACATGCTGTGTGCTG 660  
Db |||||  
Qy 923 GAGACACTCGGAGCGGCGGCCCCAGGCAAACTTGCACGAGCCCTGC 969  
Db |||||  
Qy 661 GAGACACTCGGAGCGGCGGG-CCCAGGCAAACTTGCACGAGCCCTGC 706  
Db |||||

Search completed: August 19, 2003, 08:29:08  
Job time : 3305.95 secs

**THIS PAGE BLANK (USPTO)**